National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL 339 (r. 88/04) c

Canada

THE UNIVERSITY OF ALBERTA

Discriminant Functions in Tree Breeding

by

© P.A. Jefferson

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF Doctor of Philosophy

Department of Forest Science

EDMONTON, ALBERTA

Spring 1989

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-52974-1

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR P.A. Jefferson

TITLE OF THESIS Discriminant Functions in Tree Breeding

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED Spring 1989

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) P.A. Jefferson

PERMANENT ADDRESS:

1, OAKFIELD CLOSE
WHICKHAM
ENGLAND

DATED 5TH APRIL 1989

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Discriminant Functions in Tree Breeding submitted by P.A. Jefferson in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Forest Science.

Bruce P. Dancik

Supervisor

James Beck

Robert T. Hardin

External Examiner

Date 5th April 1989

ABSTRACT

Discriminant functions which are potentially useful in tree breeding are discussed. The choice is between BLUP (Best linear unbiased predictor) and selection index. Selection indices are computationally less demanding and are favoured by tree breeders at the present time. The discriminant function should accurately predict genetic values and incorporate economic values to give a prediction of genetic worth.

There are four contributions to the development and application of discriminant functions to tree breeding:

1. In a maritime pine (*Pinus pinaster* Ait.) provnenance-progeny trial rank correlations for predicted random genetic effects using BLUP and selection index were high (.998). Caution should be used in extrapolating results to other populations and experimental designs.
2. Selection index formula are adapted to include heterogeneity among within-provenance variance-covariance matrices. Practical application of the formula requires sampling 20 to 30 families from each provenance in provenance-progeny trials.
3. Choice of economic values can reduce efficiency of selection indices. The sensitivity of the prediction variance, prediction error variance and efficiency of selection indices to changes in economic values are

assessed. Low prediction variance and relatively high genetic worth variance are associated with inefficient indices. The economic values which result in low effeciency depend on the phenotypic and genetic matrices of the breeding population.

4. Classical decision making techniques are applied to selection of economic values in the construction of selection indices for a maritime pine breeding population produced from a hierarchical mating design. Economic values for minmax and Bayes strategies are determined for density traits in the martime pine breeding population.

In the future BLUP should play a more prominent role in tree breeding because breeding populations will become more inbred and breeders will want to predict genotype performance over a range of different site types. The problems of incorporating BLUP predictions into tree breeding are discussed.

ACKNOWLEDGEMENTS

To acknowledge all the people who contributed to a project which took over three years to complete and spanned two continents is impossible. The following only recognizes major contributions by individuals and organisations. My thanks and apologies to those who assisted me and are not specifically included.

I would like to thank Dr. P. Baradat at Bordeaux for his guidance and for the opportunity to take data from his genetic trials of maritime pine. The help and guidance of Dr. B.P. Dancik is acknowledged, especially the freedom which he gave me to persue my own ideas. I am indebted to Dr. R. T. Hardin for teaching me the basics of statistics and his keen interest throughout my thesis. Thanks are due to Dr. T. Taerum for his enthusiasm and his formidible programming skills, particularly for his UANOVA program.

The assistance of Mr. M. Arbez and all his staff in the research station at Bordeaux requires acknowledgement. The technical assistance of the wood quality laboratory at Nancy is acknowledged. I am indebted to the University of Alberta and the Canadian Forest Service block grant for financial support.

A special thanks is reserved for Dr. P. Baradat and his family who had the great kindness to take a PhD. student into their home for two months. Finally I would like to

thank my parents who provided moral support throughout the project.

Table of Contents

CHAPTER ONE. INTRODUCTION 1
  I. Introduction 1
  II. Discriminant functions 2
    A. Best prediction 3
    B. Best linear prediction (selection index) 3
      i. Smith-Hazel selection index 5
        a. Multiple trait selection where information
          from related individuals is not included 5
        b. Single trait selection where information
          from related individuals is included 6
        c. Multiple trait selection where information
          from related individuals is included 7
      ii. Baradat's selection index 9
      iii. Fixed effects and selection index 10
    C. Best linear unbiased predictor (BLUP) 11
    D. Some additional discriminant functions 14
      i. Restricted discriminant functions 14
      ii. Base indices 15
    E. Variance component estimation 15
  III. Requirements of tree breeding 16
    A. Heterogeneous sites 17
    B. Heterogeneous populations 18
    C. Uncertainty over economic values 19
  IV. Research methods 20
  V. Conclusions 21
  VI. Literature cited 22
CHAPTER TWO. COMPARISON OF BLUP AND SELECTION INDEX PREDICTIONS IN A MARITIME PINE PROVENANCE PROGENY TRIAL 26
  I. Introduction 26
  II. Materials and methods 27
    A. Biological matrial 29
    B. Selection index model 29
    C. BLUP model 31
    D. Prediction of performance values 34
  III. Results and discussion 35
  IV. Conclusions 43
  V. Literature cited 45
CHAPTER THREE. SELECTION OF SUPERIOR GENOTYPES FROM SEVERAL POPULATIONS WITH HETEROGENEITY AMONG WITHIN-POPULATION VARIANCES 47
  I. Introduction 47
  II. Materials and methods 48
    A. Biological material 48
    B. Statistical techniques 48
      i. Derivation of selection index formulae 51
      ii. Validation of variance component substitution 52
        a. Calculation of the elements of the C matrix 53
        b. Calculation of the elements of the V matirx 55
      iii. Calculation of family and within-family
        variance components for each provenance 56
      iv. Selection indices 57

v. Evaluation of selection indices 57
III. Results and discussion 58
IV. Conclusion 65
V. Literature cited 67
CHAPTER FOUR. SENSITIVITY OF SELECTION INDICES TO CHANGES IN ECONOMIC VALUES 68
I. Introduction 68
II. Materials and methods 69
A. Biological material 69
B. Determination of basic components for selection index calculation 69
C. Selection index calculation 70
D. Sensitivity analysis 72
E. Calculation of index properties 73
i. Prediction variance 73
ii. Prediction error variance 73
iii. Efficiency 74
III. Results and discussion 76
A. Prediction variance response surface 77
i. Lean -10.0 circumference +10.0 81
ii. Lean -10.0 circumference -10.0 82
iii. Lean -10.0 circumference -2.00 83
B. Genetic worth variance 87
C. Standard deviation of prediction 88
D. Standard deviation of genetic worth 88
E. Prediction error variance 94
F. Efficiency 94
G. Choice of economic values 96
IV. Conclusion 97
V. Literature cited 99
CHAPTER FIVE. SELECTION STRATEGIES INCORPORATING RISK AND UNCERTAINTY OVER ECONOMIC VALUES IN TREE BREEDING .100
I. Introduction 100
II. Materials and methods 101
A. Biological material 101
B. Statistical analysis 101
C. Modification of prediction formula 105
i. Calculation of elements of V matrix 106
ii. Calculation of covariances between genetic values and phenotypic predictors 111
III. Incorporation of decision making theory into selection of economic values 112
A. Decision theory 112
B. Calculation of strategy values 115
IV. Results and discussion 118
V. Conclusion 126
VI. Literature cited 127
CHAPTER SIX. PREDICTION METHODS FOR THE FUTURE 128
I. Introduction 128
II. Future demands of tree breeding 128
III. Mixed model predictions 129
IV. BLUP versus selection index 132
V. Potential problems of implementing BLUP in tree breeding 134

A. Numerical analysis 134
B. Expected gain calculations and BLUP 136
i. Calculation of gain for decision makers 137
ii. Gain formula for selection strategies incorporating uncertaintyover economic values ...138
VI. Conclusions 139
VII. Literature cited 140
APPENDIX I. DERIVATION OF SELECTION INDEX FORMULA 142
APPENDIX II. EQUIVALENCE OF BARADAT'S INDEX TO SMITH-HAZEL SELECTION INDEX 147
APPENDIX III. DERIVATION AND CALCULATION OF BLUP PREDICTIONS 148
I. Derivation of basic BLUP formula 149
II. The equivalence of BLUP to selection indices when fixed effects are maximum likelihood estimates 150
III. Mixed model methodology to determine BLUP solutions 152
IV. Literature cited 156
APPENDIX IV. CONSTRUCTION OF BARADAT'S SELECTION INDEX FOR PROVENANCE-PROGENY MODEL 157
I. Construction of the C matrix 160
A. Covariance between the individual's additive genetic value and its provenance phenotypic value .161
B. Covariance between the individual's additive genetic value and its family phenotypic value 162
C. Covariance between the individual's additive genetic value and its within-family phenotypic value 163
II. Construction of the V matrix 164
A. Variance of provenance phenotypic value 164
B. Variance of family phenotypic value 165
C. Variance of within-family phenotypic value 166
APPENDIX V. UANOVA PROGRAM 167
I. Calculation of effects 167
II. Calculating the sum of squares from effects 172
III. Estimating variance components 174
IV. Literature cited 185
APPENDIX VI. INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER TWO. 186
I. Selection index calculations 186
A. Prediction of genetic values 192
i. Calculation of the c vector 193
ii. Calculation of the V matrix 195
iii. Prediction of genetic values from the c vector, V matrix and phenotypic predictors 197
II. BLUP calculations 197
A. Calculation of the coefficients matrix 198
B. Calculation of the matrix vector products on the right hand side of the equation 208
C. Premultiplication of the right hand side of the equation by the inverse of the coefficients matrix 208
III. Conclusion 209

APPENDIX VII. INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER THREE 211
- I. Phenotypic effects 211
- II. Variance components 230
- III. Number of trees in each family 234
- IV. Calculation of C and V matrices 234
  - A. Calculation of C and V matrices for homogeneous family variance model 235
    - i. Calculation of the elements of the C matrix 235
    - ii. Calculation of the elements of the V matrix 238
  - B. Calculation of C and V matrices for heterogeneous family variance model 241
    - i. Calculation of the elements of the C matrix 242
    - ii. Calculation of the elements of the V matrix 242
- V. Calculation of selection index values 242
- VI. Calculation of variance of predicted genetic worth, variance of genetic worth and efficiency of the selection index 244
  - A. Variance of predicted genetic worth 245
  - B. Variance of genetic worth 245
  - C. Efficiency of selection index 245
- VII. Conclusion 246

APPENDIX VIII. NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER FOUR 247
- I. Introduction 247
- II. Values plotted in chapter IV 247
- III. Verification of plotted values 251
- IV. Conclusion 251

APPENDIX IX. INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER FIVE 252
- I. Input values 252
  - A. Variance-covariance matrices 252
  - B. The number of trees in each family 253
  - C. Economic weightings 254
- II. Calculation of gain 254
  - A. C matrix for prediction of the additive genetic value of the individual due to the male parent contribution 255
    - i. The C sub-matrix between the genetic values and the male family phenotypic predictors 255
    - ii. The C sub-matrix between the genetic values and the female family phenotypic predictors 256
    - iii. The C sub-matrix between the genetic values and the within-family phenotypic predictors 257
  - B. C matrix for prediction of the additive genetic value of the individual due to the female parent contribution 258
    - i. The C sub-matrix between the genetic values and the male family phenotypic predictors 258
    - ii. The C sub-matrix between the genetic values and the female family phenotypic predictors 259
    - iii. The C sub-matrix between the genetic values and the within-family phenotypic predictors 260
  - C. V matrix for prediction of additive values 261

C. V matrix for prediction of additive genetic values 261
i. The V sub-matrix of the male family phenotypic predictors 261
ii. The V sub-matrix of the female family phenotypic predictors 262
iii. The V sub-matrix of the within-family phenotypic predictors 263
D. Calculation of $C'V^{-1}C$ matrices 264
E. Substitution of $C'V^{-1}C$ matrix in the gain formula 265
III. Calculation of min-max strategy 266
IV. Calculation of Bayes strategy 267
V. Calculation of minmax regret strategy 268
VI. Calculation of Bayes regret strategy 270
VII. Conclusion 270

List of Tables

Table I-1. List of symbols 4
Table II-1. List of symbols 28
Table II-2. Rank correaltion between BLUP and selection index rankings for linear combinations of fixed and random effects 36
Table II-3. Number of trees which rank in the top 18 of both BLUP and selection index predictions 37
Table II-4. Block, provenance, family and predicted values of the individuals ranked in the top 18 on the basis of predicted random effects 37
Table II-5. Block, provenance, family and predicted values of the trees ranked in the top 18 on the basis of predicted random effects and average value of all block effects for both BLUP and selection index 40
Table II-6. Block, provenance, family and predicted values of the trees ranked in the top 18 for BLUP but not occur in the top 18 for selection index on the basis of predicted random and fixed block effects 41
Table II-7. Block, provenance, family and predicted values of the trees ranked in the top 18 for selection index but did not occur in the top 18 for BLUP on the basis of predicted random and fixed block effects 41
Table II-8. Fixed effects block and mean estimates for BLUP and selection index 41
Table III-1. List of symbols 50
Table III-2. Provenance, family and index values of individuals in the best 40 of both indices 60
Table III-3. Provenance, family and index values of treeswhich occurred only in the selected group ranked on homogeneous index values 61
Table III-4. Provenance, family and index values of treeswhich occurred only in the selected group ranked on heterogeneous index value 62
Table III-5. Variance of predicted genetic worth, variance of genetic worth and the correlation $r(\hat{w},w)$ for each provenance 63
Table III-6. Eigen values of within-provenance variance-covariance matrices for each provenance 64
Table IV-1. List of symbols 70
Table IV-2. Range of economic weightings for circumference and lean 73
Table IV-3. Summary of index parameters 76
Table IV-4. Response surfaces and formula for calculating basic values 77
Table IV-5. $C'V^{-1}C$ matrix 79
Table IV-6. Effect of economic values on the sign of the economic value variance-covariance products 85

Table IV-7. Matrix of additive genetic variance-covariance expressed as a percentage of phenotypic variance-covariance 85
Table IV-8. G matrix 88
Table V-1. List of symbols 103
Table V-2. Variance-covariance of male family 104
Table V-3. Variance-covariance of female family 104
Table V-4. Representation of outcome values for n strategies and p states of nature 113
Table V-5. First set of economic values (minimum density held constant) 118
Table V-6. Second set of economic values (minimum density held constant) 118
Table V-7. Economic values of optimum strategies for first set of economic values 119
Table V-8. Economic values of optimum strategies for second set of economic values 119
Table AV-1. Basic data for example 168
Table AV-2. Input values for algorithm 168
Table AV-3. Summary of sum of squares and mean squares for sample data 174
Table AV-4. Summary of expected mean squares calculated by the synthesis method 183
Table AVI-1. Provenance phenotypic effects 186
Table AVI-2. Family phenotypic effects 187
Table AVI-3. Within-family phenotypic effects 188
Table AVI-4. Height variance components 192
Table AVI-5. Number of trees in each family 192
Table AVI-6. Data set for BLUP calculation example 198
Table AVI-7. BLUP estimates of effects 209
Table AVII-1. Provenance phenotypic effects 211
Table AVII-2. Family phenotypic effects 212
Table AVII-3. Within-family phenotypic effects 214
Table AVII-4. Provenance variance-covariance components .231
Table AVII-5. Family variance-covariance components for homogeneous family variance model 231
Table AVII-6. Within-family variance-covariance components for homogeneous family variance model 231
Table AVII-7. Family variance-covariance components for heterogeneous family variance model 232
Table AVII-8. Within-family variance-covariance components for heterogeneous family variance model ..233
Table AVII-9. Number of trees in each family 234
Table AVII-10. Within-provenance variance-covariance components for homogeneous family variance model 236
Table AVII-11. Families from which C and V were calculated for the estimation of index parameters ...241
Table AVIII-1. Prediction variance response surface values 248
Table AVIII-2. Genetic worth variance response surface values 248
Table AVIII-3. Prediction standard deviation response surface values 249

Table AVIII-4. Genetic worth standard deviation response surface values 249
Table AVIII-5. Prediction error response surface values .250
Table AVIII-6. Efficiency response surface values 250
Table AIX-1. Within-family variance-covariance matrix ...252
Table AIX-2. Number of trees in each family 253
Table AIX-3. Set of economic weights in numerical example (minimum density held constant). 254
Table AIX-4. Expected gain for ranges of economic values 267
Table AIX-5. Minimum outcome values for each set of economic values at selection time 267
Table AIX-6. Bayes strategy values for each set of economic values at selection time 268
Table AIX-7. Expected regret values for ranges of economic values 269
Table AIX-8. Maximum regret values for each set of economic values at selection time 269
Table AIX-9. Bayes regret strategy values for each set of economic values at selection time 270

List of Figures

Figure IV-1. Response surface of prediction variance to changes in economic values 78

Figure IV-2. Response surface of genetic worth variance to changes in economic values 86

Figure IV-3. Response surface of prediction standard deviation to changes in economic values 89

Figure IV-4. Response surface of genetic worth standard deviation to changes in economic values 90

Figure IV-5. Response surface of prediction error variance to changes in economic values 92

Figure IV-6 Response surface of correlation $R(\hat{w},w)$ to changes in economic values 93

Figure V-1. Surface of strategy values for minmax technique (first set of economic values) 120

Figure V-2. Surface of strategy values for minmax regret technique (first set of economic values) 121

Figure V-3. Surface of strategy values for minmax technique (second set of economic values) 123

Figure V-4. Surface of strategy values for Bayes technique (second set of economic values) 124

CHAPTER ONE

INTRODUCTION

I. Introduction

Breeders attempt to improve populations by selecting the 'best' individuals to be the parents of the next generation. The objective of the breeder is to select the subset of individuals which will maximize expected genetic gain. Genetic values of quantitative traits cannot be measured directly, they can only be predicted as mathematical functions of phenotypic values. These functions are known as discriminant functions because truncation selection of the population on the basis of function values will discriminate the superior genotypes. The problem has become one of finding the function which most accurately predicts the breeding value (Bulmer 1985, Henderson 1963, Henderson 1977, Searle 1974). The input values of the function are the phenotypic or record values of several traits and the economic values of the traits. The output is a linear combination of economically weighted breeding values which maximizes the economic value of the selected population.

Although discriminant functions have been used in tree breeding most applications have involved mass selection (Cotterill 1985). There are some examples where the index has incorporated information from relatives (Baradat 1976; Talbert 1984). One of the most advanced applications of selection index in tree breeding is in the maritime pine (*Pinus pinaster* Ait.) improvement program in southern France where a library of computer programs have been developed and selection index is a routine technique used in selection (Baradat 1976; Baradat 1979; Baradat 1982). Generally animal breeders have been much more sophisticated than tree breeders in their use of discriminant functions (Henderson 1973; Henderson and Quaas 1976).

It is the objective of this contribution to review the discriminant functions used in plant and animal breeding and to propose areas in which further work is required to adapt these functions to the tree breeders requirements.

II. Discriminant functions

Discriminant functions are used by staticians. A discriminant is a function which separates a population into two distinct groups. The objective being to correctly assign individuals to a group. There are several different approaches to the problem (Anderson 1984). Discriminant functions used in plant and animal breeding are based on the techniques developed by statisticians. Genetics and statistics are combined to produce functions which provide the breeder with values which can be used to assign superior genotypes to the selected population.

Discriminant functions in genetic selection can be classified into three major groups:

1. Best prediction

2. Best linear prediction (Selection index)
3. Best linear unbiased prediction (BLUP)(Henderson 1973).

There are several symbols used in this introduction. The first time a symbol is used a written explanation is given. A table of symbols is also included which gives an explanation of symbols used in the text (Table I-1).

A. Best prediction

The best predictor is the function which produces the prediction which minimizes the variance:

$$E(w - \hat{w})^2 \quad ...(I-1),$$

where w is the value to be predicted, and

$\hat{w}$ is the prediction.

The function may be linear or non-linear. Best prediction is demanding in that it requires a knowledge of the forms and parameters of the joint distribution of the phenotypic predictors and genetic values. These parameters are rarely known, as a consequence best prediction has recieved little attention from breeders.

B. Best linear prediction (selection index)

Best linear prediction (selection index) is a linear function which minimizes the value of formula I-1. Best linear prediction requires a knowledge of only the means and variances of the joint distribution of the phenotypic predictors and the genetic values. The most widely known selection index is the Smith-Hazel selection index (Smith 1937; Hazel 1943; Lin 1978).

Table I-1. List of symbols

Symbol	Explanation
w	genetic worth
$\hat{w}$	predicted genetic worth
b	phenotypic weighting
p	phenotypic value
a	economic value
v	genetic value
$\hat{v}$	predicted genetic value
$\mathbf{b}$	vector of phenotypic weightings
$\mathbf{p}$	vector of phenotypic values
$\mathbf{G}$	matrix of the genetic variance
$\mathbf{a}$	vector of economic values
$\mathbf{v}$	vector of genetic values
$\hat{\mathbf{v}}$	vector of predicted genetic values
p_i	the phenotypic value of the i^{th} individual
p_{hsfi}	the phenotypic value of the half-sib family to which the i^{th} individual belongs
p_{wfi}	the within-family phenotypic value of the i^{th} individual
$\mathbf{V}$	the phenotypic variance matrix
$\mathbf{C}$	the covariance between the phenotypic values and additive genetic values
$\hat{\mathbf{v}}$	vector of predicted genetic values
$\mathbf{X}$	design matrix of fixed effects
$\mathbf{R}$	matrix of the variance of the error effects
$\mathbf{Z}$	design matrix of random genetic effects
$\mathbf{D}$	variance of the genetic effects vector
$\boldsymbol{\beta}$	vector of unknown fixed effects
$\mathbf{u}$	vector of unknown genetic effects
$\mathbf{y}$	vector of record values
$\mathbf{A}$	numerator relationship matrix
σ_a^2	additive genetic variance

Selection indices can be of varying degrees of complexity. There are three major types:

1. Multiple trait selection where information from related individuals is not included (mass selection).

2. Single trait selection where information from related individuals is included.
3. Multiple trait selection where information from related individuals is included.

i. Smith-Hazel selection index

a. Multiple trait selection where information from related individuals is not included

The selection index is a multiple linear regression of economically weighted genetic values on weighted phenotypic values. The index value of an individual is:

$$I = b_1p_1 + b_2p_2 + \ldots + b_qp_q \quad \ldots(\text{I-2}),$$

where I is the index value of the individual,

$b_1 \ldots b_q$ are the phenotypic weightings or b values, and

$p_1 \ldots p_q$ are the phenotypic values.

The index value to be predicted is a linear combination of economically weighted genetic values for each trait. The 'b' values are calculated to maximize the genetic worth of the selected population. The genetic worth is a linear combination of genetic values:

$$w = a_1v_1 + a_2v_2 + \ldots + a_qv_q \quad \ldots(\text{I-3}),$$

where w is the genetic worth of the individual,

$a_1 \ldots a_q$ are the economic values of the q traits, and

$v_1 \ldots v_q$ are the genetic values of the q traits.

For example consider the tree breeder who wants to select for height, diameter and volume. If the economic values are correct, the 'b' values will improve the height, diameter and volume so that the economic value of future populations produced from the selected population is maximized.

b. Single trait selection where information from related individuals is included

When the breeder wants to take advantage of the genetic structure of the population it is possible to incorporate the phenotypic values of related individuals to improve the prediction of genetic values. Phenotypic values of related individuals are included by decomposing an individual's phenotypic value into linear components of phenotypic values of the genetic groups within the population. For example the phenotypic value of an individual in a population of half-sib families can be decomposed according to the linear model:

$$p_i = p_{hsfi} + p_{wfi} \quad \ldots (I-4),$$

where p_i is the phenotypic value of the i^{th} individual,

p_{hsfi} is the phenotypic value of the half-sib family to which the i^{th} individual belongs, and

p_{wfi} is the within-family phenotypic value of the i^{th} individual.

The half-sib family and within-family phenotypic values are use to predict the genetic value of the individual. The selection index for a single trait becomes;

$$I = b_1p_1 + b_2p_2 + \ldots + b_np_n \quad \ldots(I\text{-}5),$$

where $b_1 \ldots b_n$ are the b values applied to the phenotypic values of the genetic groups, and

$p_1 \ldots p_n$ are the phenotypic values of the genetic groups.

For example consider a tree produced by a heirarchical mating design. There are three phenotypic predictors:

1. The phenotypic value of the male parent family.
2. The phenotypic value of the female parent family.
3. The within family phenotypic value for the individual.

c. Multiple trait selection where information from related individuals is included.

Incorporating information on related individuals into multiple trait selection gives a new index:

$$I = b_{1,1}p_{1,1} + b_{1,2}p_{1,2} + \ldots + b_{n,q}p_{n,q} \quad \ldots(I\text{-}6),$$

where I is the index value of the individual,

$b_{1,1} \ldots b_{n,q}$ are the b values of $\underline{n}$ genetic groups for $\underline{q}$ traits, and

$p_{1,1} \ldots p_{n,q}$ are the phenotypic values of the $\underline{n}$ genetic groups for the $\underline{q}$ traits.

This is a combination of the previous two indices, for example if a breeder wants to select for height, diameter and volume in a breeding population produced by a heirarchical mating design. There will be nine 'b' values. These nine values are split into 3 sets of 3, each set corresponding to the 'b' values for the male parent family, female parent family and within family phenotypic values. Within each set the three 'b' values correspond to the height, diameter and volume traits. The 'b' values will maximize the gain in economic value.

The calculation of 'b' values in the Smith-Hazel selection index is based on the derivation of selection index formula (see Appendix I.). The 'b' values which maximize expected economic gain are given by the formula:

$$b = V^{-1}Ca \quad \ldots(I.-7)$$

where b is a (n×q) × 1 vector of b values,

V^{-1} is the inverse of a (n×q) × (n×q) matrix of the variance of the phenotypic predictors,

C is a (n×q) times q matrix of covariances between phenotypic predictors and genetic values, and

a is a q times 1 vector of economic values of n subvectors of the same q economic values.

In the previous example of multiple trait selection in a population produced by a hierarchical mating design the V matrix would be block diagonal, each block corresponding to a phenotypic predictor. The individual blocks would be 3×3 corresponding to the height, diameter and volume. The covariance matrix C is a 9×3 matirx which is subdivided into 3 submatrices each of which are of order 3×3. The 3 submatrices correspond to the covariances between:

1. The individual's genetic values and the male family

phenotypic values.

2. The individual's genetic values and the female family phenotypic values.
3. The individual's genetic values and the within family phenotypic values.

The individual elements of the submatrices correspond to the covariance between phenotypic predictor and genetic values for height, diameter and volume. The **a** vector is of order 3×1, each element being the relative economic value for height, diameter and volume.

ii. Baradat's selection index

The original concept of selection index was derived assuming 'a' values are known. Once the 'a' values had been determined 'b' values were calculated. Rouvier (1969) proposed that genetic values of individuals be predicted using information on relatives and then economic values be applied to the predictions to give the predicted genetic worth of the individual. This concept was applied to tree breeding by Baradat (1979). The prediction formula has a slightly different form for Baradat's selection index than for the Smith-Hazel selection index. The vector of genetic values of an individual is calculated by the formula:

$$\hat{\mathbf{v}} = \mathbf{C'V^{-1}p} \quad \ldots(\text{I-8}),$$

where $\hat{\mathbf{v}}$ is a $(n{\times}q) \times 1$ vector of predicted genetic values,

and

$\mathbf{p}$ is a $(n{\times}q) \times 1$ vector of phenotypic predictors.

The index value of each individual is:

$$I = \mathbf{a}'\hat{\mathbf{v}} \quad \ldots (I-9).$$

This index appears to be different from the Smith-Hazel selection index because there is no formula for the classic 'b' values. Given the same 'a' values the methods are equivalent. For proof of the equivalence of these methods see Appendix II.

iii. Fixed effects and selection index

Previously the discussion has been limited to a model of random genetic effects. In forest genetics most experiments have levels of environmental effects in addition to genetic effects. For example a progeny field trial can be laid out in a number of ways over a number of sites. Within sites, experimental designs range in complexity from randomized complete block to an incomplete latin square (Montgomery 1984). Blocks represent within site environmental effects and can be considered fixed or random. Talbert (1984) calculated a selection index for loblolly pine (*Pinus teada* L.) assuming that block effects were random. The block effects were included in the calculation of the components of the **V** matrix. Baradat (1986) considered blocks to be fixed and removed the effects before calculating the selection index. There are other examples of blocks being treated as fixed effects in forest genetics (Christophe and Birot 1979). The removal of block effects is applicable to any design, however when the breeder is faced with an incomplete block design he or she is obliged to remove the block effects. There are breeders who neglect block effects. This creates a selection index of individuals with biased random effects. There is no proof that a selection index will maximize gain in this case (Henderson 1963).

If fixed effects are included, the prediction model becomes a mixed model. Fixed effects are generally estimated in the analysis of variance and then selection index predictions are calculated from the random phenotypic effects. The predicted value is a combination of fixed effects estimates and the random predictions for the genetic values.

C. Best linear unbiased predictor(BLUP)

BLUP was developed by C.R. Henderson over a number of years. The first publication explaining the basic BLUP theory was in 1963 (Henderson 1963). Since then there have been numerous publications concerning BLUP (Henderson 1984). The theoretical basis of BLUP is based on the original approach taken by Henderson which broke away from selection index theory. Henderson concentrated on the accurate prediction of genetic values. Later Portnoy (1982) and Bulmer (1985) proved that accurate prediction maximizes gain. BLUP has the advantage of requiring only a knowledge of the variance components of the joint distribution of phenotypic and genetic values (Henderson 1973). The basic problem BLUP was developed to solve was the mixed model prediction, for example breeders need to predict random genetic effects and fixed environmental effects (Henderson 1973).

The acronym BLUP explains the basis of the BLUP derivation. The BLUP prediction is an unbiased predictor i.e. the expectation value of the records within a given fixed effect have an expectation value equal to the value of the fixed effect plus the mean. The predictor is best linear because within the class of unbiased predictors the predictor is the one which gives minimum variance (Henderson 1973). The derivation and application of BLUP estimates are based on:

1. The derivation of the basic BLUP formula.
2. The proof of the equivalence of BLUP to selection index when fixed effects are maximum likelihood estimates.
3. The use of mixed model methodology to determine BLUP solutions.

For a precise mathematical explanation of these steps the reader is refered to Appendix III.

BLUP was originally used in animal science for sire evaluation (Henderson 1973). BLUP has the advantage of estimating herd year season fixed effects from record values and a knowledge of the variance components. There were potential computational difficulties but these were avoided by considering only one trait and a relatively small number of sires.

The mixed model equations for the estimates of β and μ are (Searle 1972):

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix} \cdot \begin{bmatrix} \beta \\ u \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \quad \text{(I-10)}$$

where X is a design matrix of the fixed effects,

R^{-1} is the inverted matrix of the variance of the error effects vector,

Z is a design matrix of the random genetic effects,

D^{-1} is the inverted matrix of the variance of the genetic effects vector,

β is a vector of unknown fixed effects,

u is a vector of the unknown genetic effects, and

y is a vector of record values.

D can be decomposed into:

$$D = A\sigma_a^2 \quad \text{...(I-11)},$$

where A is the numerator relationship matrix for the individuals represented in the genetic effects vector, and

σ_a^2 is the additive genetic variance.

To illustrate some of the numerical analysis and computing problems consider the example of a population of one thousand trees. The breeder wants to predict the genetic value of 3 traits for each tree. The D and R matrix would be of order 3,000 × 3,000. There would be in excess of 3,000 unknowns in the system of equations.

For the single variable case the model can be simplified (see Appendix III). A major breakthrough in the calculation of BLUP estimates for large data sets was the development of an algorithm based on mendelian principles to invert the relationship matrix for large data sets (Henderson 1976; Quaas 1976). The use of iterative techniques to solve the BLUP equations allowed equations with up to 100,000 unknowns to be solved (Schaeffer and Kennedy 1986). These two developments coupled with increased performance of computer hardware have lead to multiple trait BLUP applications to predict breeding values of individual animals in cattle, sheep and pig breeding (Blair and Pollak 1984; Hudson and Kennedy 1985; Quaas and Pollak 1980). Despite these advances there is still a computational problem in predicting breeding values in very large breeding programs (Blair and Pollak 1984).

D. Some additional discriminant functions

i. Restricted discriminant functions

Breeders in both animal and plant breeding have used restricted disriminant functions. These functions maximize gain in genetic worth with the restriction that gain in one trait is controlled at a level specified by the breeder. There are several types of restricted selection indices (Mallard 1972). The formula for restricted BLUP has also been derived (Henderson 1984). This contribution will not consider restricted discriminant functions because these functions are a special case of a more general problem.

ii. Base indices

Williams (1962) used the term base index to refer to an index which consisted of economically weighted phenotypic values. When genetic variances cannot be accurately determined the base index can give gains which are equivalent to those for more complexed indices (Elgin *et al*. 1970). However when accurate estimates of genetic variances are available more sophisticated indices should produce greater gains.

E. Variance component estimation

The efficiency of selection indices and BLUP is dependent on the accurate estimation of variance components. There have been several studies reviewing the estimation of variance components (Searle 1974, Kennedy 1981). In an attempt to avoid repeating previous studies this series of contributions will choose one method of variance component estimation. This will allow efforts to be concentrated on the other aspects of disriminant functions used in breeding. Searle (1974) remarked on the potential of the synthesis method to estimate variance components. The major problem was the computational effort required to compute the variance components. A new sum of squares algorithm developed by Dr. T. Taerum at the University of Alberta has solved the computational problem. Kennedy (1981) found that analysis of variance methods for variance components estimation compared favourably with other estimators. Therefore the synthesis method will be used in the analysis in the subsequent series of contributions (chapters II, III, IV and V).

III. Requirements of Tree Breeding

All breeding programs follow the basic schema of selection of superior genotypes which will produce superior individuals either for the next generation of the breeding cycle or for production of a new improved crop. The genetics of the base population, reproductive physiology and crop production methods influence the prediction methods which are used to select superior genotypes. In agriculture wheat breeders are presented with a base population within which there appears to be little apparent variation. Improvement is achieved by selection among highly inbred lines. The wheat breeder can quickly produce new varieties and is relatively confident about the economic value of the final improved production crop. The relatively homogeneous cultivated fields in which wheat is grown allow the breeder to concentrate his efforts on the performance of the inbred lines in a limited number of relatively homogenous environments.

In contrast the tree breeder is generally faced with a base population which has a large amount of within and between population variation. The long production cycle, normally between 60 and 120 years in conifers, reduces the confidence in the predicted economic value of the final crop. There is a large amount of environmental variation within and between forest site types. Most forestry sites are not cultivated. The tree breeder therefore requires a prediction method which can predict the genetic value of individuals selected from heterogeneous populations grown on heterogeneous sites. In addition a prediction method must be capable of allowing for uncertainty in the economic contribution of each trait to the end product value.

A. Heterogeneous sites

The breeder requires an accurate prediction of genetic values of trees planted out in field trials where there are several levels of environmental effects. If the breeder considers environmental effects as fixed effects the prediction model is a mixed model. BLUP was derived to solve a similar problem in animal science (Henderson 1973). However the development of a BLUP program for use in a tree breeding program would require considerable effort. Computer programs for selection indices are easier to develop for the moderately large data sets analysed at the present time in tree breeding programs.

There is a need to compare BLUP results with selection index after records have been adjusted for fixed effects using analysis of variance estimates of fixed effects. If there is no practical difference between BLUP and selection index predictions the tree breeder could use the less computationally demanding selection index. At the present time most tree breeders are struggling to predict genetic values of individuals on a single site. Comparison of BLUP and selection indices results for genotypes on a single site would give a practical assessment of the potential losses of using selection indices instead of BLUP.

B. Heterogeneous populations

Accurate prediction of genetic values when variance-covariance matrices are homogeneous within a population has been proven to maximize gain in truncation selection (Portnoy 1982). Bulmer (1985) asserts that truncation selection on the basis of the most accurate prediction always maximizes gain. It is unlikely populations selected from species with large geographical ranges will have homogeneous among-population variance-covariance matrices. Populations at different ends of a geographical range would most likely have been subject to different selective forces. The gene frequencies and therefore the variance-covariance matrices would be different. There are no selection index formulae incorporating heterogeneous among-population variance-covariance matrices. There is a need to derive formulae which incorporate heterogeneous among population variance-covariance matrices. The predictions should then be compared with the predictions using homogeneous variance-covariance matrices to see if there are any practical differences.

C. Uncertainty over economic values

Once the breeder has accurately predicted genetic values he or she must determine economic values. The most certain thing in tree breeding is the uncertainty over future economic values. Baradat (1979) recognised this uncertainty and performed a sensitivity analysis of expected gain in individual traits to changes in economic values. There have been studies which simulated the effects of errors in economic weights on the efficiency of selection indices (Vandepitte and Hazel 1977). The senstitivity of the variance of predicted genetic worth, the prediction error variance and the efficiency of the index to changes in economic values have not been studied on a real breeding population. Information on this sensitivity will give the breeder additional information when choosing economic values.

Classical decision making theory provides several different techniques to take account of uncertainty (Blackwell and Girshick 1954). These techniques have never been applied to the choice of economic values in tree breeding or any genetic selection. Basic gain formulae would have to be adapted so that the gain for several possible economic values could be calculated. However once optimum strategies were developed at least they could be defended by the basic logic of decision making. It would still be up to the breeder to choose the economic values but he or she would now have additional information to aid decision making.

IV. Research methods

Prediction of genetic values or any unknown values is a comparatively new science. They are no tests such as the F test to test the significance of results (Henderson 1984). In animal science many scientists have contributed to the development of BLUP applications. The general form is an adaption of an existing formula to solve a specific problem, then an examination of results from applying the new formula to a data set. It would be logical that development of selection index formula in tree breeding would follow the same pattern. Predictions considered to be superior should be checked against predictions which are regarded as theoretically inferior. The objective should be to produce predictions which give practical differences in selected populations. There are a number of parameters which can be used to assess the quality of an index:

1. Variance in the predicted genetic worth.
2. Prediction error of the genetic worth.
3. Efficiency measured by the correlation between predicted and actual genetic worth (Henderson 1973).

Two indices can be compared using the correlation between rankings for each index and the number of individuals ranked in an upper truncated group for each index.

V. Conclusions

At the present time the most significant contribution to the application of discriminant functions to tree breeding would be to:

1. Establish the practical equivalence of BLUP and selection index when fixed effects are estimated in the analysis of variance.
2. Adapt selection index formula to take account of heterogeneous among population variances.
3. Assess the sensitivity of predicted genetic worth, prediction error variance and index efficiency to potential changes in economic values.
4. Apply classical decision making techniques to selection of economic values.

Where appropriate, new indices should be evaluated and compared against existing indices.

VI. Literature cited

Anderson, T.W. 1984 An introduction to multivariate statistical analysis, Second edition. John Wiley and Sons, New York 724 pages.

Baradat, P. 1976 Use of juvenile-mature relationships and information from relatives in combined multitrait selection. *In* Proceedings of the IUFRO joint meeting of genetic working parties on advanced generation breeding. INRA. pp. 121-138.

Baradat, P. 1979 Selection combinee multicaractere chez le pin maritime. Divers modeles d'index de selection utilizes. 104 eme Congres de Societes Savantes. Fasc. 2. pp. 299-314.

Baradat, P. 1982 Genetique quantitative modeles statistiques et genetiques de base. INRA. 204 pages.

Baradat, P. 1986 personnal communication.

Blackwell, D. and M.A. Girshick. 1954 Theory of games and statistical decisions. Dover Publications, Inc. New York. 345 pages.

Blair, H.T. and E.J. Pollak. 1984 Comparison of an animal model and equivalent reduced animal model for computational efficiency using mixed model methodology. Journal of Animal Science 58:1090-1096.

Bulmer, M.G. 1985 The mathematical theory of quantative genetics. Clarendon Press, Oxford. 254 pages.

Christophe, C. and Y. Birot. 1979 Genetic variation within and between populations of Douglas fir. Silvae Genetica 28:197-206.

Cotterill, P.P. 1985 On index selection II. Simple indices which require no genetic parameters or special expertise to contruct. Silvae Genetica 34:64-68.

Elgin, J.H., Hill R.R. and K.E. Zeiders. 1970 Comparison of four methods of multiple trait selection for five traits in alfalfa. Crop Science 10:190-193.

Hazel, L.N. 1943 The genetic basis for constructing selection indexes. Genetics 28:476-490.

Henderson, C.R. 1963 Selection index and expected genetic advance. *In* Statistical Genetics and Plant Breeding, National Research Council Publication No. 982. National Academy of Science, Washington, D.C. pp. 141-163.

Henderson, C.R. 1973 Sire evaluation and genetic trends. *In* Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. J.L. Lush. Blacksburg, Virginia. pp. 10-41.

Henderson, C.R. 1976 A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.

Henderson, C.R. 1977 Prediction of future records. *In* Proceedings of the International Conference on Quantitative Genetics. Edited by Pollak, E., O. Kempthorne and T.B. Bailey. The State University Press, Ames Iowa pp. 615-638.

Henderson, C.R. 1984 Applications of linear models in animal breeding. University of Guelph. Guelph, Ontario 462 pages.

Henderson, C.R. and R.L. Quaas. 1976 Multiple trait evaluation using relatives records. Journal of Animal Science 43:188-197.

Hudson, G.F.S. and B.W. Kennedy. 1985 Genetic evaluation of swine for growth rate and backfat thickness. Journal of Animal Science 61:83-91.

Kennedy, B.W. 1981 Variance component estimation and prediction of breeding values. Can. J. Genet. Cytol. 23:565-578.

Lin, C.Y. 1978 Index selection for genetic improvement of quantitative characters. Theor. Appl. Genet. 52:49-56.

Mallard, J. 1972 La theorie et la calcul des index de selection avec restrictions: synthesis critique. Biometrics 28:713-735.

Montgomery, D.C. 1984 Design and analysis of experiments. 2nd Edition, John Wiley and Sons, New York. 538 pages.

Portnoy, S. 1982 Maximizing the probability of correct ordering random variables using linear predictors. Journal of Multivariate Analysis 12:256-269.

Quaas, R.L. 1976 Computing the diagonal elements and the inverse of a large numerator relationship matrix. Biometrics 32:949-953.

Quaas, R.L. and E.J. Pollak. 1980 Mixed model methodology for farm and ranch beef cattle testing programs. Journal of Animal Science 51:1277-1287.

Rouvier, R. 1969 Ponderation des valeurs genotypiques dans la selection par index sur plusieurs caracters. Biometrics 25:295-308.

Schaeffer, L.R. and B.W. Kennedy. 1986 Computing strategies for solving mixed model equations. Journal of Dairy Science 69:575-579.

Searle, S.R. 1972 Linear models. John Wiley and Sons, New York. 532 pages.

Searle, S.R. 1974 Prediction, mixed models, and variance components. *In* Reliability and biometry, statistical analysis of lifelength. SIAM. Philadelphia pp. 229-266.

Smith, H.F. 1937 A discriminant function for plant selection. Ann. Eugen. 7:240-250.

Talbert, C.B. 1984 Analysis of several approaches to multiple-trait index selection in loblolly pine(*Pinus teada* L.). Phd. Thesis The University of North Carolina, Raleigh 106 pages.

Williams, J.S. 1962 The evaluation of a selection index. Biometrics 18:375-393.

Vandepitte, W.M. and L.N. Hazel. 1977 The effect of errors in economic weights on the accuracy of selection indexes. Ann. Genet. Sel. Anim. 9:87-104.

CHAPTER TWO

COMPARISON OF MIXED BLUP AND SELECTION INDEX PREDICTIONS IN A MARITIME PINE PROVENANCE-PROGENY TRIAL

1. Introduction

Traditionally, selection indices (best linear predictors) have been used to predict genetic values in tree breeding (Cotterill and Jackson 1985). Many of the linear models in tree breeding experiments are mixed models. In theory best linear unbiased predictors (BLUP) have better properties for prediction of genetic effects in mixed models (Henderson 1963). BLUP predictions are equivalent to selection index predictions when random genetic values are predicted from phenotypic values which have been calculated from record values adjusted by maximum likelihood estimates of fixed effects (Gianola and Goffinet 1982). The computational time required to produce BLUP estimates is much greater than that required to calculate selection index values (Baker 1986; Schaeffer and Kennedy 1986). To produce maximum likelihood estimates of fixed effects to ensure that selection index values are exactly the same as BLUP estimates is also computationally demanding (Henderson 1973). To estimate fixed effects in the analysis of variance and then calculate selection indices from adjusted record values is much less demanding. Tree breeders have continued to use selection indices as opposed to BLUP because selection indices are computationally less demanding. In theory selection index predictions are biased (Henderson

1963). The practical differences between BLUP and selection index predictions in tree breeding populations have not yet been investigated.

It is the objective of this contribution to compare the rankings of individuals in a maritime pine (*Pinus pinaster* Ait.) provenance-progeny trial ranked on the basis of BLUP and selection index with fixed effects estimated in the analysis of variance.

II. Materials and methods

There are no computer programs available for the BLUP model to predict individual genotypes in a provenance-progeny trial. A BLUP model was developed in order to make an accurate comparison between BLUP and selection index. To avoid many of the numerical analysis problems involved in the calculation of BLUP estimates for large data sets a small data set was choosen and the analysis was limited to a single trait. The objective being to obtain BLUP equations which could be solved by inversion of the coefficients matrix. There are several symbols used in this contribution. The first time a symbol is used a written explanation is given. A table of symbols is also included which gives an explanation of symbols used in the text.

Table II-1. List of symbols

Symbol	Explanation
Y_{lijk}	record value of the k^{th} tree nested within the j^{th} family in the i^{th} provenance in the l^{th} block
μ	mean for all records in the analysis
$BLOCK_l$	Block effect of the l^{th} block
$PROV_i$	Provenance effect of the i^{th} provenance
$FAM_{j(i)}$	Phenotypic family effect of the j^{th} family in the i^{th} provenance
$e_{k(ji)}$	Within family effect of the k^{th} individual in the j^{th} family
$\hat{v}$	predicted additive value
I	Index value
c	vector of covariance between phenotypic predictors and the additive genetic value
V	matrix of variance of phenotypic predictors
p	vector of phenotypic predictors
y	vector of record values
X	design matrix of fixed effects
β	vector of unknown fixed effects
Z	design matrix of random effects
a	vector of unknown random genetic effects
eb	vector of residual effects
Z_1	design matrix of random provenance additive genetic effects
a_1	vector of unknown provenance additive genetic effects
Z_2	design matrix of random within provenance additive effects
a_2	vector of unknown within provenance additive genetic effects
G_1	numerator relationship matrix for provenances
G_2	numerator relationship for individuals
s_1	ratio of residual variance to the provenance additive genetic variance
s_2	ratio of residual variance to the within-provenance additive genetic variance
σ^2_{eb}	variance of residual effect in BLUP model
σ^2_e	within-family variance
σ^2_a	within-provenance additive genetic variance
σ^2_{GP}	variance of provenance genetic effects
σ^2_{PROV}	provenance variance
σ^2_{FAM}	family variance

Table II-1. List of symbols (continued)

Symbol	Explanation
V_B	variance-covariance of residual random effects vector of the BLUP model after fixed effects have been removed
R	variance-covariance matrix of the residual effects in the BLUP model

A. Biological material

The basic data was taken from a 12-year-old provenance-progeny trail of maritime pine. The complete trial comprised 795 individuals from 68 half-sib families taken from 12 provenances planted in three replications with five incomplete blocks in each replication. Forty families from seven provenances in three incomplete blocks of the experiment were used for the comparison of BLUP and selection index predictions. Each tree was measured for height. To provide accurate variance component estimates the variance components were estimated from the complete data set. Block effects were considered fixed in both BLUP and selection index predictions.

B. Selection index model

The basic linear model in the analysis of variance is:

$$Y_{lijk} = \mu + BLOCK_l + PROV_i + FAM_{j(i)} + e_{k(ji)} \quad \ldots(II-1)$$

where Y_{lijk} is the record value (height) of the k^{th} tree nested within the j^{th} family of the i^{th} provenance in the l^{th} block;

μ is the mean for all records in the analysis;

$BLOCK_l$ is the effect of the l^{th} block;

$PROV_i$ is the phenotypic effect of the i^{th} provenance;

$FAM_{j(i)}$ is the phenotypic family effect of the j^{th} family nested in the i^{th} provenance, and

$e_{k(ji)}$ is the phenotypic effect of the k^{th} individual nested within the j^{th} family in the i^{th} provenance.

The fixed effects and the random phenotypic effects were estimated from the linear model by the UANOVA analysis of variance program developed by Dr T. Taerum at the University of Alberta (Appendix V). Variance components were estimated in the same program (Appendix V).

The selection index calculation was the one pioneered by Baradat (1982). The index value is a linear combination of economically weighted predicted additive genetic values. In the case of single trait selection no economic weightings are required. In the provenance-progeny trial model the additive genetic value of an individual is the sum of two independent additive values. The two additive values are:

1. The provenance additive value.
2. The within-provenance additive value.

The prediction formula for the additive genetic value of each individual is:

$$\hat{v} = c'V^{-1}p \quad ...(II-2),$$

where $\hat{v}$ is the predicted additive value;

c is a vector of the covariance between the phenotypic predictors and the additive genetic value,

V^{-1} is the inverse of the variance matrix of the phenotypic predictors, and

p is a vector of the estimated phenotypic values.

The individual's provenance and within-provenance additive values were predicted using equation II-2. The total additive value of the individual or in this case the selection index value was given by adding these two predictions. The c and V^{-1} matrices are determined from basic genetic and statistical principles (Appendix IV). A numerical example of the selection index calculation is given in Appendix VI.

C. BLUP model

BLUP models were originally developed for sire evaluation in animal science. Later models were developed to evaluate individual animals (Henderson and Quaas 1976). The individual animal model will be used here because it is equivalent to the selection index model where genetic values of individuals are predicted. The BLUP model is based on the basic linear model:

$$y = X\beta + Za + eb \quad ...(II-3),$$

where y is a vector of record values,

$\mathbf{X}$ is a design matrix of fixed effects for each record,

$\boldsymbol{\beta}$ is a vector of unknown fixed effects,

$\mathbf{Z}$ is a design matrix of random effects,

$\mathbf{a}$ is a vector of unknown genetic effects, and

$\mathbf{eb}$ is a vector of residual effects.

The model requires modification to include the two sets of random additive genetic values in the provenance-progeny model i.e. provenance additive value and within-provenance additive value. The modified linear model is:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{a}_1 + \mathbf{Z}_2\mathbf{a}_2 + \mathbf{eb} \; \ldots (\mathrm{II}\text{-}4),$$

where $\mathbf{y}$ is a vector of record values,

$\mathbf{X}$ is a design matrix of fixed effects for each record;

$\boldsymbol{\beta}$ is a vector of unknown block effects,

$\mathbf{Z}_1$ is a design matrix of random provenance additive genetic effects,

$\mathbf{Z}_2$ is a design matrix of random within-provenance additive genetic effects,

$\mathbf{a}_1$ is a vector of unknown provenance additive genetic effects,

$\mathbf{a}_2$ is a vector of within-provenance additive genetic effects, and

$\mathbf{eb}$ is a vector of residual effects.

The provenance and within-provenance additive genetic effects are considered to be independent of each other. Using mixed model methodology to set up the BLUP equation gives (Henderson 1977):

$$\begin{bmatrix} X'X & X'Z_1 & X'Z_2 \\ Z_1'X & Z_1'Z_1 + G_1^{-1}s_1 & Z_1'Z_2 \\ Z_2'X & Z_2'Z_1 & Z_2'Z_2 + G_2^{-1}s_2 \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = \begin{bmatrix} X'y \\ Z_1'y \\ Z_2'y \end{bmatrix} \quad \text{(II-5)}$$

where G_1^{-1} is the inverse of the numerator relationship matrix for the provenances,

G_2^{-1} is the inverse of the numerator relationship matrix for individuals,

s_1 is the ratio of the residual variance to the provenance additive genetic variance, and

s_2 is the ratio of the residual variance to the within-provenance or individual additive genetic variance.

G_1 is an identity matrix. G_2 is a block diagonal matrix. The blocks of the block diagonal matrices are of order equal to the number of individuals within a family. The diagonal elements of the blocks are all equal to one. The off diagonal elements are equal to 0.25. The s_1 ratio is:

$$s_1 = \frac{\sigma_{eb}^2}{\sigma_{GP}^2} \quad \ldots\text{(II-6)}$$

where σ_{eb}^2 is the variance of the residual effects in the linear model (II-4), and

σ_{GP}^2 is the variance of the provenance genetic effects.

The term σ_{eb}^2 is calculated from the variance components in the analysis of variance and is equal to:

$$\sigma_{eb}^2 = \sigma_e^2 - 3\sigma_{FAM}^2 \quad \text{...(II-7)},$$

where σ_e^2 is the variance of the residual calculated in the analysis of variance, and

σ_{FAM}^2 is the family variance calculated in the analysis of variance.

The variance σ_{GP}^2 is by definition equal to the component σ_{PROV}^2 calculated in the analysis of variance. The s_2 ratio is:

$$s_2 = \frac{\sigma_{eb}^2}{\sigma_a^2} \quad \text{...(II-8)},$$

where σ_a^2 is the within provenance additive genetic variance.

The term σ_{eb}^2 is calculated in the same way as for s_1. The variance σ_a^2 was calculated from the variance components in the analysis of variance and is equal to four times σ_{FAM}^2. BLUP estimates were calculated by premultiplying the vector on the right hand side of the mixed model equations (II-5) by the inverse of the coefficient matrix. A numerical example of the BLUP calculation is given in Appendix VI.

D. Prediction of performance values

Breeders want to predict the performance of genotypes. If the genotypes are measured in blocks or replications there are a number of linear combinations which the breeder can use to assess performance. The breeder may be interested in the performance of the genotypes when the fixed block effects have been removed. Breeders may be interested in including fixed block effects or an average value for all the fixed block effects in predictions (Henderson 1973).

In this study each individual was given a BLUP estimated score and a combination of selection index and analysis of variance estimates for:

1. Random provenance additive genetic effect plus random within-provenance additive genetic effect.
2. Fixed block effect plus random provenance additive genetic effect plus random within-provenance additive genetic effect.
3. Average fixed block effect plus random provenance additive genetic effect plus random within-provenance additive genetic effect.

The performance of each tree in the breeding population was predicted by BLUP and a combination of selection index scores and fixed effects estimated in the analysis of variance. Trees were ranked on predicted performance values and the rank correlation between the BLUP and the selection index score rankings was calculated. The number of trees which occurred in the top 18 (top ten percent) for both BLUP and selection index scores was determined.

III. Results and discussion

The information obtained can be used to compare the prediction methods. The rank correlations and the trees in the selected population provide information for estimating the practical differences between prediction methods. If the rank correlation is high and the top 18 ranked trees are the same for both prediction methods the breeder is free to choose the computationally less demanding method. Comparison of predicted values and an examination of the elements of each prediction formula will give insights into situations where prediction methods may differ significantly.

The rank correlations and the trees selected in the top 18 ranked on the basis of predicted additive genetic effects for both BLUP and selection index predictions show that if the breeder is only interested in random additive genetic effects the selected population is the same for both prediction methods (Tables II-2 and II-3). The populations selected when the linear combination of predicted average block effects plus the predicted random additive genetic effects is used as a performamance measure are identical for both prediction methods (Tables II-2 and II-3). When a linear combination of block plus random additive genetic effects is predicted there are substantial differences between the selected populations for each prediction method. (Table II-3).

Table II-2 Rank correlation between BLUP and selection index rankings for linear combinations of fixed and random effects

Random	Block + random	Block(ave) + random
.998	.910	.998

Table II-3 Number of trees which ranked in the top 18 individuals of both BLUP and selection index predictions

Random	Block + random	Block(ave) + random
18	9	18

The rank correlations for performance based on random genetic effects are high but the predicted performance values of the trees differ slightly (Table II-4).

Table II-4 Block, provenance, family and predicted values of the trees ranked in the top 18 on the basis of predicted random effects for both BLUP and selection index predictions

	Identification		Prediction	
Block	Provenance	Family	BLUP	Index
2	3	4	271.6	264.8
2	3	4	225.2	218.3
2	3	4	190.3	183.5
1	1	2	188.2	189.0
3	6	2	167.1	172.1
2	4	2	159.0	160.1
3	5	3	158.3	155.8
3	7	3	153.4	145.1
1	1	2	147.5	148.3
1	1	2	147.5	148.3
2	3	4	143.8	137.0
1	3	1	136.6	144.3
2	5	3	132.4	135.4
3	6	2	132.3	137.3
3	6	3	130.5	135.4
3	6	2	126.5	131.1
1	1	2	124.3	125.1
3	5	3	123.4	121.0

The difference between the two predicted values can be traced to the way in which the block effects are estimated. In the selection index method block effects are estimated and then record values are adjusted. The genetic effects are then predicted from the adjusted record values. This is a two stage process. In BLUP the fixed and random maximum likelihood estimates are calculated at the same time. The two stage process of adjusting records for fixed effects and then calculating selection index predictions will give exactly the same results as BLUP if the fixed effects are maximum likelihood estimates (Gianola and Goffinet 1982). Thus the difference between random predictions must be due to the different block estimates for BLUP and selection index calculations.

The maximum likelihood estimates for fixed effects are given by the formula:

$$\beta = (X'V^{-1}X)^{-}(X'V^{-1}y) \quad \ldots (II\text{-}9).$$

The V matrix in the maximum likelihood estimate is equal to:

$$V = ZGZ' + R \quad \ldots (II\text{-}10),$$

where Z is the design matrix of random genetic effects,

G is the variance covariance matrix of the random genetic effects vector, and

R is a diagonal matrix of residual effects (Henderson 1984).

The V matrix in genetic experiments is not always diagonal because the G matrix can have several covariance terms. The covariance terms are the covariances between related individuals. The analysis of variance method assumes that V is a diagonal matrix. Each diagonal term is a linear combination of the error variance and the random effect variance as calculated in the analysis of variance.

The similarity of the predicted random genetic effects estimates indicate a certain robustness of random predictions to different fixed effects estimates. In certain populations the V matrix may depart significantly from a diagonal matrix. For example in a population which is highly inbred and has low effective population size there would be a large number of related individuals. It is likely that BLUP and selection index predictions of random genetic effects would diverge in such circumstances.

There is a marked difference between the BLUP and selection index prediction values for the linear combination of average block effect plus random additive genetic effects (Table II-5), however rankings for both prediction methods are the same (Table II-2). An explanation can be found in the method of calculation. BLUP estimates for block effects are based on unadjusted raw record values. The block estimates in the analysis of variance are calculated from deviations from the overall mean. The BLUP estimate is therefore much larger because it includes the mean. The rankings are not changed because the average of the block estimates is a constant. Adding a constant to the values will not change the rank correlation. It will be the same as the rank correlation when trees are ranked according to predicted random additive genetic values.

Table II-5 Block, provenance, family and predicted values of the trees ranked in the top 18 on the basis of predicted random effects and an average value of all block effects for both BLUP and selection index predictions

	Identification		Prediction	
Block	Provenance	Family	BLUP	Index
2	3	4	942.6	265.5
2	3	4	896.2	219.1
2	3	4	861.3	184.2
1	1	2	859.2	189.8
3	6	2	838.1	172.9
2	4	2	830.0	160.9
3	5	3	829.3	156.6
3	7	3	824.4	145.8
1	1	2	818.5	149.1
1	1	2	818.5	149.1
2	3	4	814.9	137.8
1	3	1	807.6	145.1
2	5	3	803.4	136.1
3	6	2	803.3	138.0
3	6	3	801.5	136.5
3	6	2	797.5	132.2
1	1	2	795.3	125.9
3	5	3	794.2	121.8

When the objective is to predict a linear combination of fixed block and random genetic effects there are substantial differences in predicted values and rankings. An examination of the block number of the trees which were found only in the top 18 of the selection index rankings and those found only in the top 18 of the BLUP rankings showed that the block estimates are the major factor reducing the rank correlations (Tables II-6 and II-7).

Table II-6 Block, provenance, family and predicted values of the trees ranked in the top 18 for BLUP but which did not occur in the top 18 for selection index on the basis of predicted random and fixed block effects

	Identification		Prediction	
Block	Provenance	Family	BLUP	Index
1	3	1	840.7	120.5
1	1	4	840.6	113.1
1	3	4	838.0	115.5
1	3	1	834.9	114.7
1	1	5	830.5	103.7
1	3	4	826.4	103.9
1	2	3	824.3	93.6
1	2	5	820.9	90.5
1	3	1	817.4	97.2

Table II-7 Block, provenance, family and predicted values of the trees ranked in the top 18 for selection index but which did not occur in the top 18 for BLUP on the basis of predicted random and fixed block effects

	Identification		Prediction	
Block	Provenance	Family	BLUP	Index
3	6	2	806.1	178.5
3	5	3	797.2	162.2
3	7	3	792.3	151.5
2	4	2	782.5	133.4
3	6	2	771.2	143.7
3	6	3	769.4	142.1
3	6	2	765.4	137.9
3	5	3	762.3	127.4
3	6	2	753.8	126.3

Table II-8 Fixed effects block and mean estimates for BLUP and selection index.

Estimate	Mean	Block 1	Block 2	Block 3
BLUP	-	750.6	623.5	638.9
Index	673.3	22.6	-26.7	6.4

The differences between BLUP estimates for block 1 and the other blocks are greater than the same differences for block effects estimated in the analysis of variance (Table II-8). This difference explains the influx of individuals from block 1 into the top 18 ranked BLUP predictions. When selection index predictions are used, individuals from Block 3 replace individuals from Block 1 in the top 18 ranked individuals (Tables II-6 and II-7). The BLUP block estimates are maximum likelihood estimates and the analysis of variance estimates are least squares estimates. These estimates do effect rankings when they are directly included in the prediction. In theory the maximum likelihood estimate is better because it give an unbiased estimate. The V^{-1} matrix takes into account the relationships between the genetic groups. The constraints provided to solve equations may also have an effect on the estimates. If the breeder has the computational capabilities he or she would be best advised to use the BLUP fixed effects estimates.

The results show that if the breeder wants to predict random additive genetic effects when fixed effects have been removed, analysis of variance estimates of fixed effects give almost identical rankings as the BLUP estimates. The rank correlation may be lower in other populations where there are many related individuals and the V matrix departs significantly from a diagonal matrix. Considering the extra computational effort required for BLUP estimates the breeder can reduce the computional problems and obtain the same selected population by using selection index and analysis of variance estimates. If a population has a large number of related individuals selection index predictions should be viewed with some caution.

The results are based on data from one experiment. These results are dependent on experimental data and experimental design. Some caution should be exercised in extrapolating the results to other situations. When the relative magnitude of fixed and random effects are different from those in the experiment results may not be applicable. Analytical methods exist which would be more appropriate for extrapolation of results to other data sets particularly with respect to BLUP estimations (Henderson 1975).

IV. Conclusions

In the provenance-progeny study analysed in this experiment results show breeders wanting to include fixed effects in their predictions should use BLUP estimates. If the breeder wants to rank individuals on the basis of random genetic effects after removal of fixed effects selection index with records adjusted using analysis of variance estimates for fixed effects will give the same results as BLUP. The breeder can use selection index, reducing the computational demands, without loosing any efficiency in selection.

For more highly inbred populations or populations with low effective numbers the breeder is advised to use BLUP estimates. Caution should be used in extrapolating the results to other data sets.

V. Literature cited

Baker, R.J. 1986 Selection indices in plant breeding. CRC. Press, Inc. Boca Raton, Florida. 218 pages.

Baradat, P. 1982 Genetique quantitative modeles statistiques et genetiques de base. INRA. 204 pages.

Cotterill, P. and N. Jackson 1985 On index selection. I. Methods of determining economic weight. Silvae Genetica 34:57-61.

Gianola D. and B. Goffinet 1982 Sire evaluation with best linear unbiased predictors Biometrics 38:1085-1088.

Henderson, C.R. 1963 Selection index and expected genetic advance. *In* Statistical Genetics and Plant Breeding, National Research Council Publication No. 982. National Academy of Science, Washington, D.C. pp. 141-163.

Henderson, C.R. 1973 Sire evaluation and genetic trends. Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. J.L. Lush. Blacksburg, Virginia. pp. 10-41.

Henderson, C.R. 1975 Comparison of alternative sire evaluation methods. Journal of Animal Science 41:760-770.

Henderson, C.R. 1977 Prediction of future records. *In* Proceedings of the International Conference on Quantitative Genetics. Edited by Pollak, E., O. Kempthorne and T.B. Bailey. The State University Press. Ames Iowa pp. 615-638.

Henderson, C.R. 1984 Applications of linear models in animal breeding. University of Guelph. Guelph, Ontario 462 pages.

Henderson, C.R. and R.L. Quaas. 1976 Multiple trait evaluation using relatives records. Journal of Animal Science 43:188-197.

Schaeffer, L.R. and B.W. Kennedy. 1986 Computing strategies for solving mixed model equations. Journal of Dairy Science 69:575-579.

CHAPTER THREE

SELECTION OF SUPERIOR GENOTYPES FROM SEVERAL POPULATIONS WITH HETEROGENEITY AMONG WITHIN-POPULATION VARIANCES

I. Introduction

When the breeder wants to predict genetic values to select the best genotypes from several different provenances there is an assumption made that the genetic and phenotypic variance-covariance matrices are all the same (Searle 1974). In statistical terminology within-provenance variance-covariance matrices are assumed to be homogeneous among populations. When provenances are sampled from species that have wide geographical ranges there is little to justify the above assumption. The discriminant function that most effectively selects superior genotypes is the index which most accurately predicts breeding value (Bulmer 1985, Henderson 1977). The selection index constructed by Baradat (1979) is based on this principle but assumes that within-provenance variance-covariance matrices are homogeneous among populations. Adaption of prediction formulae to incorporate heterogeneous variance-covariance matrices would in theory provide a more accurate prediction and therefore a more efficient selection index.

The objective of this contribution is to derive the prediction formulae when there is heterogeneity among within-provenance variance-covariance matrices. The rankings of predicted genotypes will be compared with the rankings of predicted genotypes when the within-provenance variance-covariance matrices are assumed to be homogeneous among populations.

II. Materials and methods

A. Biological material

A 12-year-old provenance-progeny trial of maritime pine (*Pinus pinaster* Ait.) containing a total of 68 half-sib families from 12 provenances was measured for height, circumference and stem lean. The 12 provenances were sampled from the species range in south-west France. Six half-sib families were sampled in each provenance, except for provenance 10 in which only five families were sampled and provenance 12 in which only three families were sampled. Seven hundred and ninety five trees from the 68 half-sib families were planted out in three replications. Each replication was subdivided into five incomplete blocks, making a total of 15 incomplete blocks.

B. Statistical techniques

The basic input values for selection index calculations are:

1. The variance components of random genetic and phenotypic effects.
2. Estimates of random phenotypic effects.

The linear model for each record value includes fixed replication and block effects. To obtain the basic input values for selection index calculations the independent variables were standardized and the fixed effects removed. The adjusted records followed the linear model:

$$Y_{ijk} = \mu + PROV_i + FAM_{j(i)} + e_{k(ji)} \quad ...(III-1)$$

where Y_{ijk} is the adjusted record value for the ijkth individual,

μ is the mean value of all records,

$PROV_i$ is the phenotypic effect of the ith provenance,

$FAM_{j(i)}$ is the phenotypic family effect of jth family nested within the ith provenance, and

$e_{k(ji)}$ is the within-family phenotypic effect of the kth individual nested within the jth family.

All variance-covariance components and phenotypic effects were estimated using the UANOVA program developed by Dr. T. Taerum at the University of Alberta (Appendix V.).

There are several symbols used to explain the derivation and calculation of selection index values. The first time a symbol is used a written explanation is given. A table of symbols is also included which gives an explanation of each symbol used in the text (Table III-1).

Table III-1. List of symbols

Symbol	Explanation
μ	mean of all records
$PROV_i$	phenotypic effect of the i^{th} provenance
$FAM_{j(i)}$	phenotypic family effect of the j^{th} family nested within the i^{th} provenance
$e_{k(ji)}$	within-family phenotypic effect of the k^{th} individual nested within the j^{th} family
$\hat{PROV}_i$	estimated phenotypic effect of the i^{th} provenance
$\hat{FAM}_{j(i)}$	estimated phenotypic family effect of the j^{th} family nested with the i^{th} provenance
$\hat{e}_{k(ji)}$	estimated within-family phenotypic effect of the k^{th} individual nested within the j^{th} family
$\bar{Y}_{...}$	average of all records
$\bar{Y}_{i..}$	average of all record values in the i^{th} provenance
$\bar{Y}_{ij.}$	average of all record values in the j^{th} family in the i^{th} provenance
Y_{ijk}	the record value for the ijk^{th} individual
$n_{i..}$	number of individuals in the i^{th} provenance
$n_{ij.}$	number of individuals in the j^{th} family in the i^{th} provenance
nf	number of families in a provenance
a_{ijk}	additive genetic value of the ijk^{th} individual
p	vector of phenotypic predictors
$\hat{v}$	vector of predicted genetic values
C	matrix of the covariance among the phenotypic predictors and genetic values
V	matrix of the variance among the pehnotypic predictors
G	matrix of the genetic variance
a	vector of economic values
VA_i	additive genetic variance-covariance in the i^{th} provenance
$\sigma^2_{FAM(PROV_i)}$	family variance-covariance for the i^{th} provenance genetic worth
$\sigma^2_{e_{(i)}}$	within-family variance for the i^{th} provenance
w	genetic worth
$\hat{w}$	predicted genetic worth
I	selection index value

i. Derivation of selection index formulae

The selection index calculated was the one proposed by Baradat(1979):

$$I = a'C'V^{-1}p \ldots(III-2),$$

where I is the index of predicted genetic worth,

a is a vector of economic values,

C is the matrix of the covariance between the genetic value and the phenotypic predictors,

V^{-1} is the inverted matrix of the variance of the phenotypic predictors, and

p is a vector of phenotypic predictors.

When the variance-covariance matrices are considered to be homogenous the elements of the C and V matrices can be calculated (Appendix IV.). The calculation of the elements of the C and V matrices changes as the number of individuals in a family and the number of families within a provenance changes. The variance-covariance components calculated in the analysis of variance, from which the elements of the C and V matrices are calculated, are pooled estimates that remain unchanged from provenance to provenance. The inclusion of heterogeneous variance in the prediction formulae requires:

1. Proof of the validity of the substitution of the pooled family and within-family variance-covariance matrices for all provenances by the family and

within-family variance-covariance matrix for each provenance in the prediction formulae.

2. Calculation of the family and within-family variance components for each provenance.

ii. Validation of variance component substitution

The substitution of pooled within-provenance variance-covariance component estimates by separate provenance variance-covariance component estimates for each provenance will be validated if it can be shown that the elements of the C and V^{-1} matrices within a provenance can be determined without reference to the within-provenance components of the other provenances. The components of the selection index formula are derived from the linear model (III-1). The phenotypic predictors are estimated from the basic records:

$$\hat{PROV}_i = \bar{Y}_{i..} - \bar{Y}_{...} \quad \text{...(III-3)},$$

$$\hat{FAM}_{j(i)} = \bar{Y}_{ij.} - \bar{Y}_{i..} \quad \text{...(III-4)},$$

$$\hat{e}_{k(ji)} = Y_{ijk} - \bar{Y}_{ij.} \quad \text{...(III-5)},$$

where $\hat{PROV}_i$ is the estimate of the phenotypic effect of the i^{th} provenance,

$\hat{FAM}_{j(i)}$ is the estimate of the phenotypic effect of the j^{th} family nested within the i^{th} provenance,

$\hat{e}_{k(ji)}$ is the estimate of the phenotypic effect of the k^{th} individual nested within the j^{th} family in the

i^{th} provenance,

$\bar{Y}_{...}$ is the average of all the record values,

$\bar{Y}_{i..}$ is the average of all the record values in the i^{th} provenance,

$\bar{Y}_{ij.}$ is the average of all the record values in the j^{th} family in the i^{th} provenance, and

Y_{ijk} is the record value for the ijk^{th} individual.

It is important to note that the estimated phenotypic predictors within a provenance are calculated from within-provenance record values and the average of all the records. Provided the grand mean is an unbiased estimate all the phenotypic predictors in any given provenance are independent of within-provenance record values in the other provenances.

a. Calculation of the elements of the C matrix

The C matrix is the covariance matrix between an individual's phenotypic predictors and an individual's additive value. The additive genetic value of the individual is split into two parts:

1. The provenance additive genetic value.
2. The within-provenance additive value.

The expectation value of the covariance between the individual's estimated phenotypic predictors and its provenance additive value is by definition equal to the provenance variance component. This is true even if variance-covariance matrices are heterogeneous among provenances.

The expectation value of the covariance between the individual's estimated phenotypic predictors and its within-provenance additive genetic value is derived by determining which individuals contribute to the phenotypic predictors. The covariance between each of the individual phenotypic values that contribute to the estimate of each predictor and the individual's additive genetic value is then expressed in terms of the within-provenance additive genetic variance (Appendix IV.). The individual phenotypic values are estimated from the within-provenance record values and the overall mean. Given that the overall mean is a constant for the whole population, the phenotypic predictors of the individuals within a provenance are independent of the within-provenance deviations of the other provenances. Therefore, all the covariances between an individual's phenotypic predictors and its within-provenance additive value can be expressed in terms of within-provenance additive variance for that given provenance. The substitution of the within-provenance variance-covariance terms in the formula for the elements of the C submatrices gives (Appendix IV.):

$$\mathrm{Cov}(a_{ijk},\hat{\mathrm{PROV}}_i) = \frac{(n_{ij.} - 1)(1 - .25)}{n_{ij.}} \mathrm{VA}_i \quad \ldots\text{(III-6)},$$

$$\mathrm{Cov}(a_{ijk},\hat{\mathrm{FAM}}_{j(i)}) = [.25(n_{ij.} - 1) + 1] \left[\frac{1}{n_{ij.}} - \frac{1}{n_{i..}}\right] \mathrm{VA}_i \,(\mathrm{III}\text{-}7),$$

$$\mathrm{Cov}(a_{ijk},\hat{e}_{k(ji)}) = \frac{.25(n_{ij.} - 1) + 1}{n_{i..}} \mathrm{VA}_i \;\ldots(\mathrm{III}\text{-}8),$$

where $n_{i..}$ is the number of individuals in the i^{th} provenance,

$n_{ij.}$ is the number of individuals in the j^{th} family in the i^{th} provenance,

a_{ijk} is the additive genetic value of the ijk^{th} individual, and

VA_i is the additive genetic variance-covariance as calculated from the family variance-covariance component in the i^{th} provenance.

b. Calculation of the elements of the V matrix

Given that the mean is an unbiased estimate, the expectation of the variance of the phenotypic values can be expressed in terms of provenance and within-provenance variance components. The derivation of the equations is the same as the calculation of expectation values of the sum of squares in the analysis of variance (Searle 1972). The elements of the V submatrices when heterogeneity among within-provenance variance components is included are:

$$\sigma_{\mathrm{PR\hat{O}V}}^2 = \frac{\left[n_{i..}^2\sigma_{\mathrm{PROV}}^2 + \sum_{j=1}^{nf} n_{ij.}^2\sigma_{\mathrm{FAM(PROV}_i)}^2\right]}{n_{i..}^2} + \frac{1}{n_{i..}}\sigma_{e(i)}^2 \;\ldots(\mathrm{III}\text{-}9),$$

$$\sigma_{\hat{FAM}}^{2} = \frac{\left[n_{i..}^{2} - \sum_{j=1}^{nf} n_{ij.}^{2}\right]}{n_{i..}^{2}} \sigma_{FAM(PROV_i)}^{2} + \left[\frac{1}{n_{ij.}} - \frac{1}{n_{i..}}\right] \sigma_{e(i)}^{2} \text{(III-10)}.$$

$$\sigma_{\hat{e}}^{2} = \frac{n_{ij.} - 1}{n_{ij.}} \sigma_{e(i)}^{2} \;...\text{(III-11)},$$

where nf is the number of families in the i^{th} provenance,

$\sigma_{FAM(PROV_i)}^{2}$ is the family variance-covariance for the i^{th} provenance, and

$\sigma_{e(i)}^{2}$ is the within family variance for the i^{th} provenance.

iii. Calculation of family and within-family variance components for each provenance

The hierarchical linear model in the analysis of variance is suited to the calculation of separate within-provenance variance components for each provenance. The analysis of variance can be performed separately on each provenance. The grand mean in the analysis of variance when all the provenances are included is different from the mean when only one of the provenances is considered. The mean when only one provenance is included is equal to the grand mean when all the provenances are included plus the provenance effect for that provenance. The expectation of the mean squares of the within-provenance effects for any given provenance will be the same in both analyses. The analysis of variances of the individual provenances will provide an estimate of the within-provenance variance-covariance components for each provenance based on the expected mean squares of the within-provenance deviations for that provenance. The individual within-provenance estimates will vary because of sampling variation. The pooled estimates based on several provenances will be less sensitive to sampling variation.

iv.Selection indices

Selection index scores or predicted genetic worth values were calculated for each individual. Two indices were calculated:

1. Assuming that within-provenance variance-covariance components were homogeneous among populations.
2. Adaption of prediction formulae to include heterogeneity among within-provenance variance-covariance matrices.

Height circumference and lean were all given an economic value of 1.0.

v. Evaluation of selection indices

The two indices scores on the martime pine breeding population were compared using:

1. Rank correlation.
2. The individuals selected at 5% selection intensity.

The index that incorporates heterogeneity among within-provenance variance-covariance matrices has a number of different C and V matrices. Each provenance has different C and V matrices. Changing the C and V matrices will alter the prediction variance, genetic worth variance and the efficiency of the index. The three index properties were calculated for each provenance to give an idea of quality of prediction within each provenance.

1. Variance of predicted genetic worth was calculated by the formula:

$$a'C'V^{-1}Ca \quad ...(III-12),$$

2. Variance of genetic worth was calculated by the formula:

$$a'Ga \quad ...(III-13),$$

   where G is a matrix of genetic variance-covariance.

3. Efficiency as measured by the correlation between genetic worth and predicted genetic worth calculated by the formula:

$$r(\hat{w},w) = (a'C'V^{-1}Ca)^{0.5}\ (a'Ga)^{-0.5} \quad ...(III-14),$$

   where $\hat{w}$ is the predicted genetic worth; and
   w is the genetic worth.

Input values and numerical examples of calculations for selection index, variance of predicted genetic worth, variance of genetic worth and efficiency are given in Appendix VII.

III. Results and discussion

The weighted average of all the within-provenance variance-covariance components was equal to the variance-covariance components when homogeneity was assumed. This confirmed the theory of the calculation of separate within-provenance variance-covariance components.

The rank correlation between the two indices was 0.66. The composition of the selected population (best 40 trees) changed from the index based on homogeneous variances to the index based on heterogeneous variances. Three major groups could be identified:

1. A group of trees which occurred in the selected population for both homogeneous and heterogeneous selection index rankings. These trees came from provenances 1, 4, 6, 8 and 11 (Table III-2).
2. A group of trees which occurred only in the selected population for homogeneous selection index rankings. These trees came from six provenances. Over half these trees come from provenance 2 (Table III-3).
3. A group of trees which occurred only in the selected population for heterogeneous selection index rankings. The majority of these trees came from provenances 1 and 12 (Table III-4).

Heterogeneous index values were higher than the homogeneous index values (Tables III-3 and III-4).

Table III-2 Provenance, families and index values of individuals in the best 40 of both indices.

Individual identification		Index values	
Provenance	Family	Homogeneous	Heterogeneous
1	2	2.19	6.01
8	4	2.12	4.42
8	4	2.06	4.85
1	2	1.88	3.89
11	3	1.86	4.11
11	2	1.83	3.04
11	2	1.83	2.42
1	2	1.79	3.29
11	3	1.72	3.09
1	2	1.72	4.90
6	2	1.71	2.55
11	3	1.65	3.24
4	2	1.55	2.30
11	3	1.55	2.83
1	2	1.52	3.50
11	3	1.51	3.08

Examination of the variance of the predicted genetic worth, variance of genetic worth and correlation $r(\hat{w},w)$ showed:

1. High variances of predicted genetic worth and genetic worth are associated with provenances 1 and 12 (Table III-5). The correlation $r(\hat{w},w)$is greater than one in these two provenances (Table III-5).
2. A negative genetic worth variance is associated with provenance 2 (Table III-5). The correlation $r(\hat{w},w)$ was calculated but it is undefined because the square root of a negative number is undefined (Table III-5).

Table III-3 Provenance, families and index values of trees which occurred only in the selected group ranked on homogeneous index values.

Individual identification		Index values	
Provenance	Family	Homogeneous	Heterogeneous
9	4	2.81	0.36
9	4	2.19	0.62
7	4	2.12	0.32
2	2	2.11	0.60
2	2	2.08	-0.35
10	3	2.04	1.25
2	3	1.97	-0.19
2	3	1.86	0.08
9	6	1.86	0.51
2	2	1.81	-0.46
10	1	1.81	1.91
9	4	1.78	0.55
2	3	1.72	0.12
2	3	1.70	0.26
2	4	1.69	0.58
3	6	1.68	1.24
7	4	1.67	-0.50
2	3	1.66	0.17
4	3	1.65	1.40
2	2	1.63	0.24
7	4	1.61	0.53
2	1	1.53	-0.95
2	4	1.53	-0.54
10	3	1.52	1.50

High correlation $r(\hat{w},w)$ values of provenances 1 and 12 are associated with high heterogeneous selection index values. However, provenance 10 has a high correlation value but the heterogeneous index values of the trees from provenance 10 are not high enough for them to be ranked in the selected population (Tables III-2 and III-4). The factor which causes a high correlation is the difference between the variance of the genetic worth and the variance of the predicted genetic worth (Table III-5). There are two ways to

Table III-4 Provenance, family and index values of trees which occurred only in the selected group ranked on heterogeneous index values.

Individual identification		Index values	
Provenance	Family	Homogeneous	Heterogeneous
12	3	0.46	8.39
12	2	1.35	6.24
1	2	1.24	5.88
12	2	1.16	5.57
1	4	0.63	4.95
12	1	0.90	4.80
12	3	-0.42	4.70
12	2	1.05	4.57
12	3	-0.44	4.35
12	2	0.92	4.26
12	1	1.04	4.11
12	2	0.96	3.54
11	5	1.37	3.45
1	2	0.96	3.12
12	1	0.89	2.97
1	1	1.23	2.82
12	3	-0.46	2.71
1	5	0.36	2.69
1	2	1.18	2.67
12	3	-0.60	2.59
6	1	0.50	2.44
1	1	1.22	2.43
6	6	1.15	2.41
8	1	1.12	2.30

obtain large differences between these values:

1. A combination of average predicted genetic worth variance and low genetic worth variance.
2. A combination of high predicted genetic worth variance and average genetic worth variance.

The variance of predicted genetic worth in provenance 10 is average. The cause of the high correlation is the low genetic worth variance (Table III-5). The high correlation value of the trees in provenances 1 and 12 are due to a combination of average genetic worth variance value and high prediction variance values. Results show that high prediction variance values in the heterogeneous index are associated with trees that are selected in the heterogeneous index but not selected with the homogeneous index.

Table III-5 Variance of predicted genetic worth, variance of genetic worth and the correlation $r(\hat{w},w)$ for each provenance

Provenance	Variance predicted genetic worth	Variance genetic worth	Correlation $r(\hat{w},w)$
1	6.93	2.69	1.60
2	0.22	-0.29	0.87
3	0.52	0.52	1.00
4	0.81	1.29	0.80
5	0.29	0.59	0.70
6	1.66	1.67	1.00
7	0.49	0.51	0.98
8	3.67	2.98	1.11
9	0.17	0.46	0.61
10	1.92	0.38	2.26
11	3.56	3.42	1.02
12	19.35	7.56	1.60

These results can be explained by sampling variation. Sampling variations of between group components in the multivariate case can cause non-positive definite variance-covariance matrices (Hill and Thompson 1978). This indicates the presence of negative variance estimates. The effects of these negative estimates produce an unstable index (Hill and Thompson 1978). There are a number of ways to estimate variance components. Some estimators are more appropriate than others depending on the experimental design (Kennedy 1981). Changing the method of variance component estimation in this example will have negligible effect. Sampling more families within each provenance and use of an appropriate variance component estimator would stablize the index.

The table of the eigen values of each variance-covariance matrix for within-provenance variance components of each provenance show a number of negative eigen values (Table III-6).

Table III-6 Eigen values of within-provenance variance-covariance matrices for each provenance

Provenance	Family			Within family		
	Eig-1	Eig-2	Eig-3	Eig-1	Eig-2	Eig-3
1	-0.05	0.01	0.61	0.10	0.70	0.93
2	-0.06	-0.03	0.05	0.14	0.83	1.45
3	-0.03	0.01	0.11	0.16	0.56	1.34
4	0.00	0.06	0.27	0.11	0.71	1.82
5	-0.01	0.05	0.11	0.12	0.74	1.46
6	0.00	0.13	0.23	0.09	0.58	1.23
7	-0.09	0.01	0.08	0.13	0.78	1.32
8	-0.06	0.00	0.33	0.14	0.77	1.28
9	0.00	0.02	0.24	0.13	0.56	1.79
10	-0.08	0.01	0.18	0.11	0.87	1.17
11	-0.10	0.00	0.42	0.11	0.76	1.94
12	-0.10	0.03	1.04	0.13	1.04	1.43

All the negative values are for family variance-covariance matrices. None of the within-family matrices have negative eigen values (Table III-6). The family variance-covariance matrices in provenances 1 and 12 both have one high positive eigen value, one low positive eigen value and one low negative eigen value. This eigen value combination tends to be associated with inflated index scores (Tables III-4 and III-6). A possible cause in provenance 12 is the low number of families in the provenance, however this cannot explain the situation in provenance 1. Provenance 2 is unique in having two low negative eigen values and one positive eigen value. This eigen value combination tends to be associated with low index scores (Tables III-3 and III-6). The non-positive definite family variance-covariance matrices affect the $C'V^{-1}$ and $C'V^{-1}C$ matrices. The affect is dependent on the eigen value combination. This is seen in the high variance of predicted genetic worth and the high index value scores for the individuals from provenances 1 and 12 which enter the selected group when heterogeneous variance-covariance components replace the homogeneous variance-covariance components in the selection index (Tables III-3 and III-4).

It is not possible to propose any rules for eigen value combinations which could be used to assess the stablity of future indices. The breeder should look at the eigen values and be suspicious of the index if any of these values are negative. High prediction variance or negative genetic worth variance will confirm that the phenotypic and genetic variances have not been accurately estimated. The index can only be improved if a greater number of families in each provenance are sampled.

IV. Conclusion

The selection index formula can be modified to take account of heterogeneity of within-provenance variance-covariance matrices among provenances. The individual within provenance family variance-covariance estimates are subject to sampling variations. If the breeder wants to predict additive genetic values taking into account heterogeneous among provenance variance-covariance terms. the number of families must be greater than six. Christophe and Birot (1979) suggested that in excess of 20 families per provenance should be tested in a second sampling after initial experiments had identified those provenances of interest. The results of this contribution confirm this strategy for breeders who want to predict the genetic worth of individuals from provenances where it is suspected that there is heterogeneity among within-provenance variance-covariance matrices.

V. Literature cited

Baradat, P. 1979 Selection combinee multicaractere chez le pin maritime. Divers modeles d'index de selection utilizes. 104 eme Congres de Societes Savantes. Fasc. 2. pp. 299-314.

Bulmer, M.G. 1985 The mathematical theory of quantative genetics. Clarendon Press, Oxford. 254 pages.

Christophe, C. and Y. Birot. 1979 Genetic variation within and between populations of Douglas fir. Silvae Genetica 28:197-206.

Henderson, C.R. 1977 Prediction of future records. *In* Proceedings of the International Conference on Quantitative Genetics. Edited by Pollak, E., O. Kempthorne and T.B. Bailey. The State University Press. Ames, Iowa pp. 615-638.

Hill, W.G. and R. Thompson 1978 Probabilities of non-positive definite between group or genetic covariance matrices. Biometrics 34:429-439.

Kennedy, B.W. 1981 Variance component estimation and prediction of breeding values. Can. J. Genet. Cytol. 23:565-578.

Searle, R.S. 1972 Linear models. John Wiley and Sons, New York. 532 pages.

Searle, S.R. 1974 Prediction, mixed models, and variance components. *In* Reliability and Biometry, Statistical Analysis of Lifelength. SIAM. Philadelhpia. pp. 229-266.

CHAPTER FOUR

SENSITIVITY OF SELECTION INDICES TO CHANGES IN ECONOMIC VALUES

I. Introduction

Selection index theory is based on the premise that genetic, phenotypic and economic parameters are known. The effects of sampling variation in genetic and phenotypic parameters has been studied extensively (Harris 1964; Hayes and Hill 1980). Results (Harris 1964) indicate 500 to 1000 individuals are required for the selection index to be reasonably efficient. In tree breeding experiments there are generally more than 1000 trees in a single experiment. Tree breeders are generally more concerned about uncertainty over economic values of traits because of the long rotation age of tree crops.

A selection index is a prediction based on economically weighted genetic values (genetic worth). There has been some work on the effects of sampling errors in the determination of economic values on the efficiency of an index (Smith 1983; Vandepitte and Hazel 1977). The effects of varying economic values on expected gain have been determined by sensitivity analysis (Baradat 1979). There has been no work on the sensitivity of the prediction variance, prediction error variance and efficiency to potential changes in economic values.

It is the objective of this contribution to assess the sensitivity of the index properties:

1. Prediction variance
2. Prediction error variance
3. Efficiency

to changes in economic values.

II. Materials and methods

A. Biological material

A 12-year-old provenance-progeny trail of maritime pine (*Pinus pinaster* Ait.) was measured for height, circumference and stem lean. There were a total of 795 trees from 68 half-sib families collected from 12 provenances in south west France. The trees were planted out in three replications, each replication subdivided into five incomplete blocks, making a total of 15 incomplete blocks.

B. Determination of basic components for selection index calculation

The basic input values for the selection index calculations are estimates of random phenotypic effects and variance estimates of the random phenotypic effects calculated in the analyses of variance. The dependent variables were standardized before the effects were calculated in the analyses of variance. The linear model for each record contains a fixed replication and block effect. Each record value was adjusted to remove the fixed block and replication effects. Random effects were calculated using the UANOVA analysis of variance program developed at the University of Alberta by Dr. T. Taerum (appendix V.). Variance and covariance components were calculated in the UANOVA program (appendix V.).

C. Selection index calculation

The selection index was the one used by Baradat (1979). There are a number of symbols (Table IV-1) required to explain the calculation of index values and the properties of an index. The first time a symbol is written an explanation is provided. In addition an explanation of all the symbols found in this chapter is given in Table IV-1.

Table IV-1. List of symbols

Symbol	Explanation
p	vector of phenotypic predictors
$\hat{v}$	vector of predicted genetic values
C	matrix of the covariance among the phenotypic predictors and genetic values
V	matrix of the variance among the phenotypic predictors
G	matrix of the genetic variance
a	vector of economic values
$\sigma_{\hat{w}}^2$	variance of predicted genetic worth
σ_w^2	variance of the genetic worth
w	genetic worth
$\hat{w}$	predicted genetic worth
i	the selection intensity

In the provenance-progeny trial model the predicted additive genetic value of an individual is the sum of two independent predicted additive values:

1. The within-provenance additive value.
2. The provenance additive value.

Because each of the predictions are independent they can be summed to give the total additive value of an individual. The index values or predicted genetic worth values are given by the linear combination:

$$\hat{w} = a'\hat{v} \ldots \text{(IV-1)},$$

where $\hat{w}$ is the predicted genetic worth or index value,

a is a vector of economic values for each trait, and

$\hat{v}$ is a vector of predicted genetic values.

The genetic values are predicted from the formula:

$$\hat{v} = C'V^{-1}p \ldots \text{(IV-2)},$$

where C is the covariance matrix between the phenotypic predictors and the genetic values,

V is the variance of the phenotypic predictors, and

p is a vector of phenotypic predictors.

The construction of the C and V matrices is given in appendix IV. For the provenance model there are three phenotypic predictors:

1. Provenance phenotypic value.
2. Family within provenance phenotypic value.
3. Within-family phenotypic value.

If three traits are incorporated into the selection p is a 9 × 1 vector, $\hat{v}$ is a 3 × 1 vector, a is a 3 × 1 vector and $\hat{w}$ is a scalar.

D. Sensitivity analysis

The index values change as the economic value vector a changes (see equation(IV-1)). There is great uncertainty over the economic value of a given trait. Baradat (1979) developed a sensitivity analysis to assess the effects of changing economic values on expected gain values. The sensitivity procedure used by Baradat was employed here to test the effects of changing economic values on prediction variance, prediction error variance and efficiency. The economic value for height was held constant at 1.0. The assumption of holding the economic value of height to 1.0 is restrictive as the index may have better properties in some cases if height was allowed to take a negative value. Increasing height is so important in most improvement programs it is difficult to argue that it should have a negative value.

A range of economic values for circumference and lean were used to produce a total of 121 different indices based on the same predicted genetic values. The range of circumference and lean economic values are given in Table IV-2. The prediction variance of genetic worth, the prediction error variance of genetic worth and the efficiency of the index were calculated for each of the 121 indices.

Table IV-2. Range of economic weightings for circumference and lean

Circumference	-10	-8	-6	-4	-2	1	2	4	6	8	10
Lean	-10	-8	-6	-4	-2	1	2	4	6	8	10

E. Calculation of index properties

i. Prediction variance

The prediction variance is the variance of the predicted genetic worth. It is derived from the regression of genetic values on phenotypic values (Rao 1965):

$$\sigma_{\hat{g}}^2 = a'C'V^{-1}Ca \quad \text{...(IV-3)},$$

where $\sigma_{\hat{g}}^2$ is the prediction variance,

a is the vector of economic values,

C is matrix of the covariance between the genetic values and the phenotypic predictors, and

V is the matrix of the variance of the phenotypic predictors.

The variance of the predicted within-provenance plus provenance additive values is the sum of $a'C'V^{-1}Ca$ for each predicted value (Kendall and Stuart 1958). In the remainder of this contribution $\sigma_{\hat{g}}^2$ will refer to the variance of the predicted within-provenance plus provenance additive values.

ii Prediction error variance

The prediction error variance is the expectation value of the squared difference between the value to be predicted and the prediction i.e.:

$$E(w - \hat{w})^2 \quad ...(IV-4),$$

where w is the genetic worth, and

$\hat{w}$ is the estimated genetic worth.

The prediction error variance can be calculated from experimental data by the formula:

$$\sigma_{(w - \hat{w})}^{2} = a'Ga - a'C'V^{-1}Ca \quad ...(IV-5),$$

where $\sigma_{(w - \hat{w})}^{2}$ is the prediction error variance, and

G is the genetic variance matrix (Henderson 1973).

The prediction error variance given by the formula above is the difference between the variance of the genetic worth and the variance of the predicted genetic worth.

iii. Efficiency

The efficiency of an index is defined as the ratio of expected gain to truncation selection on the basis of predicted (index) values ($\hat{w}$) to the expected gain from truncation selection on the basis of actual values (w) (Bulmer 1985). The expected gain using the predicted values is:

$$i\sigma_{\hat{w}} \quad ...(IV-6),$$

where i is the selection intensity; and

$\sigma_{\hat{w}}$ is the standard deviation of the predicted values.

The expected gain using actual values is:

$$i\sigma_w \ldots(\text{IV-7}),$$

where σ_w is the standard deviation of the genetic worth values.

The ratio between the two gains gives the relative efficiency:

$$\text{Relative efficiency} = \frac{i\sigma_{\hat{w}}}{i\sigma_w} \ldots(\text{IV-8}).$$

The above ratio is equal to the correlation between the predicted and actual genetic worth values (Bulmer 1985):

$$R(\hat{w},w) \ldots(\text{IV-9}).$$

The relative efficiency is calculated from equation (IV-8). Rewriting equation (IV-8) and expressing the genetic variance and predicted genetic variance in terms of the a, C and V matrices gives:

$$\frac{(a'C'V^{-1}Ca)^{.5}}{(a'Ga)^{.5}} \ldots(\text{IV-10}).$$

The efficiency formula above is the ratio of the standard deviation of genetic worth to the standard deviation of predicted genetic worth.

A summary table of the three index parameters which are used to assess the sensitivity of the index to changes in economic values is given in Table IV-3. Numerical examples of the calculation of parameter values and the values plotted in the figures in this chapter are given in appendix VIII.

Table IV-3. Summary of index parameters

Property	Expectation value	Formula
Prediction variance	$(\hat{w})^2$	$a'C'V^{-1}Ca$
Prediction error variance	$(\hat{w} - w)^2$	$a'Ga - a'C'V^{-1}Ca$
Efficiency	$R(\hat{w},w)$	$\frac{(a'C'V^{-1}Ca)^{.5}}{(a'Ga)^{.5}}$

The sensitivity of the index parameters was determined by changing the values of the a vector so that each property was calculated for the 121 combinations of 'a' values in table IV-2.

III. Results and discussion

To interpret the results of the sensitivity analysis 6 response surfaces were produced, one each for the basic values from which the properties were calculated and one for each of the properties (Table IV-4).

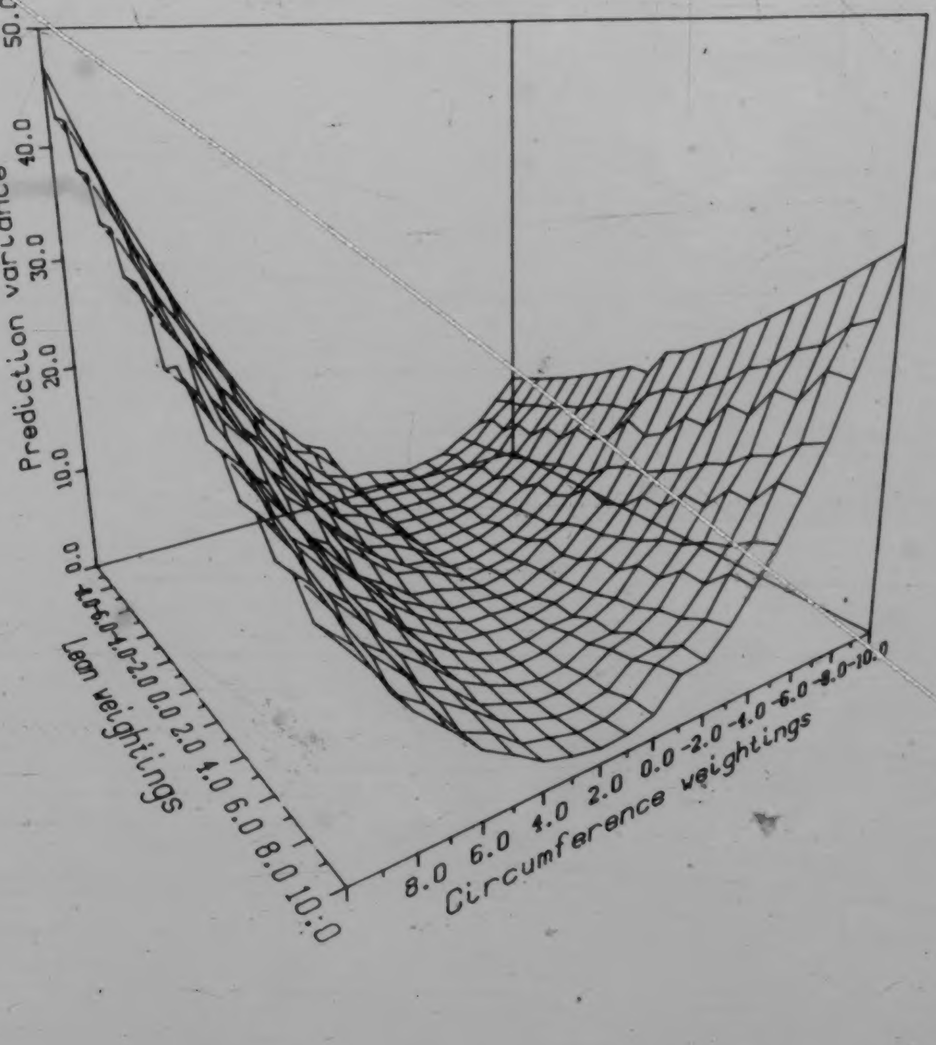
Table IV-4 Response surfaces and formula for calculating basic values

Response surface	Formula
Prediction variance	$a'C'V^{-1}Ca$
Genetic worth variance	$a'Ga$
Prediction standard deviation	$(a'C'V^{-1}Ca)^{.5}$
Genetic worth standard deviation	$(a'Ga)^{.5}$
Prediction error variance	$a'Ga - a'C'V^{-1}Ca$
Efficiency	$\frac{(a'C'V^{-1}Ca)^{.5}}{(a'Ga)^{.5}}$

A. Prediction variance response surface

The shape of the response surface is like a river valley (Figure IV-1.). The bottom of the river valley runs from the economic value coordinates, circumference -4.0 lean -10.0, to the coordinates circumference 2.0 lean 10.0 (Figure IV-1). The steepness of the valley walls is not equal. The steepest slope on the valley sides is between the economic value coordinates, circumference -4.0 lean -10.0, and coordinates, circumference 10.0 lean -10.0 (Figure IV-1). The shallowest valley side is between the economic value coordinates, circumference -4.0 lean -10.0, and cordinates, circumference -10.0 lean -10.0 (Figure IV-1.).

Figure IV-1. Response surface of prediction variance to changes in economic values



The response surface shown in Figure IV-1 is explained by the structure of the $C'V^{-1}C$ matrix (Table IV-5).

Table IV-5. $C'V^{-1}C$ matrix

	Height	Circumference	Lean
Height	0.331	0.262	-0.078
Circumference	0.262	0.230	-0.067
Lean	-0.078	-0.067	0.030

Increasing the economic value of a trait in the **a** vector will increase the influence of the $C'V^{-1}C$ components associated with that trait on the scalar produced by the vector matrix mutiplication $a'C'V^{-1}Ca$. The influence of any single trait on the prediction variance is dependent on:

1. The economic value of the trait and the economic values of the other traits in the **a** vector.
2. The magnitude of the variance and covariance components associated with the trait in the $C'V^{-1}C$ matrix.

To clarify the explanation, consider **X** to be the matrix produced by the $C'V^{-1}C$ product with the individual elements $x_{11}, x_{12}, \ldots, x_{33}$ where the subscripts refer to traits 1 to 3 (1 = height, 2 = circumference, 3 = lean). The scalar value of the prediction variance is the result of the addition of 9 scalar products produced in the vector matrix multiplication $a'Xa$. The prediction variance is equal to:

$$a_1a_1x_{11} + a_1a_2\,x_{12} + \ldots + a_3a_3x_{33} \quad \ldots\text{(IV-11)}.$$

where $a_1 \ldots a_3$ are the economic values of traits 1...3,

x_{11} is the variance component associated with trait 1, and

x_{12} is the covariance component associated with trait 1 and 2 in the $C'V^{-1}C$ matrix.

Each of the nine scalars is the product of economic value products and a variance or covariance component in the $C'V^{-1}C$ matrix. There are three economic value variance products and six economic value covariance products. The six products between economic values and covariance components are three identical pairs of products because the $C'V^{-1}C$ matrix is symetrical. The traits associated with the three basic economic value and covariance products are:

1. height-lean
2. circumference-lean
3. height-circumference

Two points should be noted:

1. The variance components in the $C'V^{-1}C$ are always multiplied by the square of the economic value of the trait associated with the variance component. Increasing the absolute value of the economic value of a trait will increase the contribution of the economic value variance product.
2. The covariance components in the $C'V^{-1}C$ are always multiplied by the product of the economic values of the traits in the covariance component. The product of economic values and covariance components may be negative if the economic value cross product is of

different sign to the covariance component.

The interaction of the economic values and the $C'V^{-1}C$ components to produce the prediction variance value will be demonstrated by explaining how the a values in the a vector and the components in the $C'V^{-1}C$ (Table IV-5) interact to produce the prediction variance at the economic value coordinates:

1. circumference +10.0
   lean -10.0
2. circumference -10.0
   lean -10.0
3. circumference -2.0
   lean -10.0

i. circumference +10.0 lean -10.0

The products between the economic values of the variance components are all positive. Using the **X** matrix terminology and substituting 'a' values into formula IV-11 the scalars from each of the economic value variance products are:

$$1x_{11} \quad 100x_{22} \quad 100x_{33} \quad \ldots\text{(IV-12)}.$$

The products between the economic values and the covariance between height and lean are positive. To improve the explanation the symbol xv will be used to express the absolute value of the elements in the **X** matrix. Substituting the 'a' values into formula IV-11 gives the economic covariance product $10xv_{13}$. The products between the economic values and the covariance between circumference and lean are positive. Substituting the 'a' values gives an economic value covariance product $100xv_{23}$. The products between the economic values and the covariance between height and circumference are positive. Substituting the 'a' values gives an economic value covariance product $10x_{12}$.

Adding all the economic value variance-covariance products gives the prediction variance:

$$x_{11} + 100x_{22} + 100x_{33} + 2(10x_{12}) + 2(10xv_{13}) + 2(100xv_{23}) \quad \ldots(\text{IV-13}).$$

Economic value products and variance-covariance components have combined to give positive values for all the scalars. The 'a' values are the most extreme values considered in this analysis. Therefore the prediction variance is maximum.

ii. circumference -10.0 lean -10.0

The scalars from each economic value variance product are the same as in the first example i.e.

$$1x_{11} \quad 100x_{22} \quad 100x_{33} \quad \ldots(\text{IV-14}).$$

The products between the economic values and the covariance between height and lean are positive. Substituting the 'a' values gives the economic value covariance product $10xv_{13}$. The products between the economic values and the covariance between circumference and lean are negative. Substituting the 'a' values gives an economic value covariance product $-100xv_{23}$. The products between the economic values and the covariance between height and circumference are negative. Substituting the 'a' values gives an economic value covariance product $-10x_{12}$.

Adding all the economic value variance-covariance products gives the prediction variance:

$$x_{11} + 100x_{22} + 100x_{33} + 2(-10x_{12}) + 2(-100xv_{23}) + 2(10xv_{13}) \quad \text{(IV-15).}$$

The negative values of the economic value covariance products for height and circumference, and circumference and lean, reduce the maximum prediction variance in comparison to the first example (Figure IV-1).

iii. circumference -2.0 lean -10.0

The products between the economic values and the variance components are:

$$1x_{11} \ 4x_{22} \ 100x_{33} \ \text{...(IV-16).}$$

The products between the economic values and the covariance between height and lean are positive. Substituting the 'a' values gives the economic value covariance product $10xv_{13}$. The products between the economic values and the covariance between circumference and lean are negative. Substituting the 'a' values gives the economic value covariance product $-20xv_{23}$. The products between the economic values and the covariance between height and circumference are negative. Substituting the 'a' values gives the economic value covariance product $-2x_{12}$.

Adding all the economic value covariance products gives the prediction variance:

$$x_{11} + 4x_{22} + 100x_{33} + 2(-2x_{12}) + 2(10xv_{13}) + 2(-20xv_{23}) \quad \text{(IV-17)}.$$

Given the relative magnitude of the variance-covariance components the components tend to cancel each other out and produce a low prediction variance value (Figure IV-1).

The gradient between two points on the response surface is dependent on the sign of the economic value coordinates of the points. The effect of the sign of the economic value of lean and circumference on the sign of the product of the economic value and covariance is given in Table IV-6.

The gradient is dependent on which quadrant of the surface the slope is located i.e. the sign combinations of the economic values. If the sign of the products of the economic values and covariance is negative and the products for the other components are positive, changing the 'a' values will not give a steep gradient. When all the economic value covariance products are positive the gradient will be steep. This explains the steep gradient in the quadrant where the economic value of lean is negative and the economic value of circumference is positive. Within a given quadrant the steepness of a slope depends on the direction of the slope with respect to the economic value axis. Slopes parallel to the lean economic value axis will slope more gradually than slopes parallel to the circumference economic value axis because of the greater magnitude of the circumference components in the $C'V^{-1}C$ matrix.

Table IV-6. Effect of economic values on sign of economic value variance-covariance product

Economic values		Product of economic values and covariance components		
lean	circumference	lean height covariance	lean circumference covariance	height circumference covariance
+	+	-	-	+
+	-	-	+	-
-	+	+	+	+
-	-	+	-	-

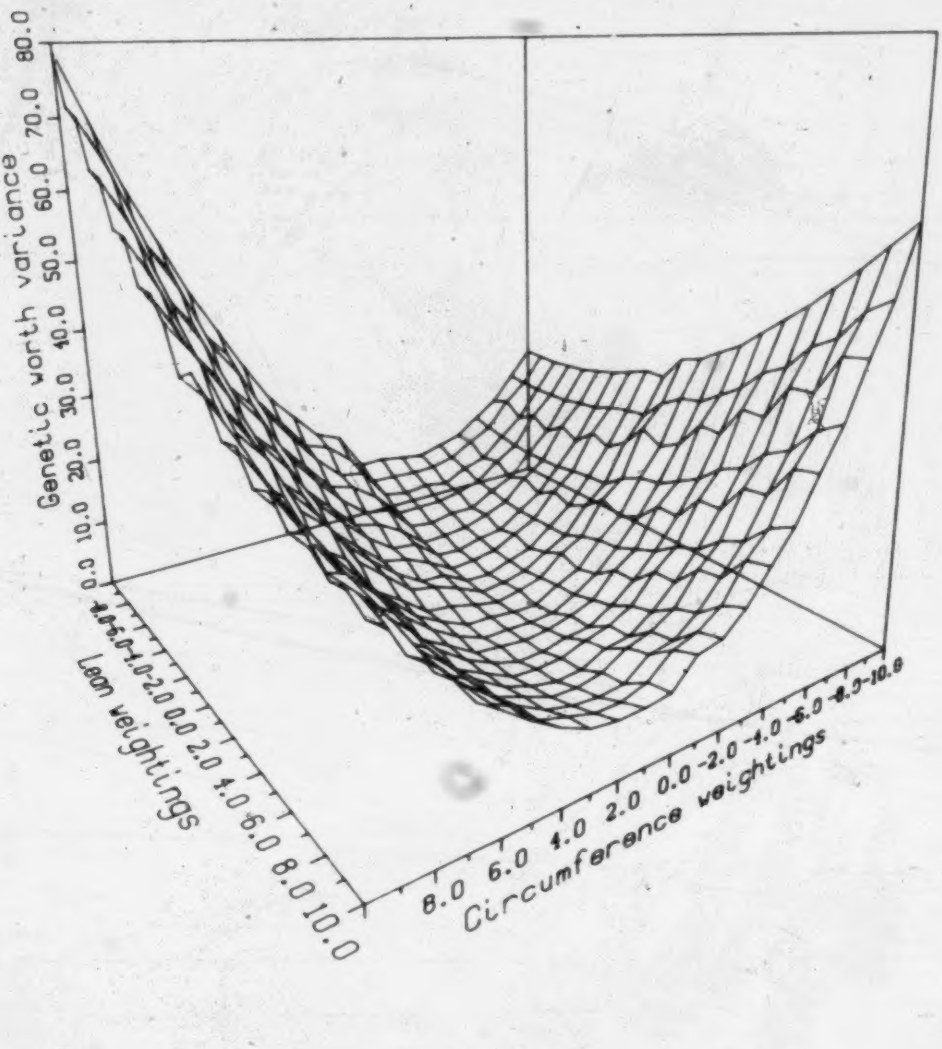
The major reason for the small lean components is the heritability of lean (0.11). The matrices of the additive genetic variance-covariance components expressed as a percentage of phenotypic variance-covariance components shows that lean is under very weak genetic control compared to height and circumference (Table IV-7).

Table IV-7. Matrix of additive genetic variance-covariance expressed as a percentage of phenotypic variance-covariance.

	Height	Circumference	Lean
Height	65.55	61.72	65.08
Circumference	61.72	51.08	36.84
Lean	65.08	36.84	10.98

The low values for lean in the $C'V^{-1}C$ matrix is a direct

Figure IV-2. Response surface of genetic worth variance to changes in economic values



result of this poor genetic control. The components of the V^{-1} are not affected by the degree of genetic control. However the components of the C matrix will be smaller if the traits associated with the components have low heritabilities.

B. Genetic worth variance

The general shape of the genetic worth variance response surface is the same as the prediction variance surface (Figure IV-2). The similarity in the shape is a result of the proportions of the individual variance-covariance components of the $C'V^{-1}C$ (Table IV-5) and G (Table IV-8) matrices being very similar in both matrices. For example the largest component in both matrices is the height variance. The lean variance component is the lowest in both matrices. There are some minor differences in the proportions of the two matrices. The relative magnitude of the covariance between circumference and lean changes slightly. In the G matrix the covariance between circumference and lean is roughly equal to the lean variance component (Table IV-8). In the $C'V^{-1}C$ matrix the covariance between circumference and lean is approximately double the lean variance component (Table IV-5). This change has no major effect on the response surface because of the low value of the circumference-lean covariance components.

The gradients of the genetic worth surface are steeper (Figure IV-2). The reason for this is that the components of the G matrix are larger than the components of the $C'V^{-1}C$ matrix (Tables IV-5 and IV-8). Thus any increase in an 'a' value will increase the genetic worth variance more than the prediction variance.

Table IV-8. G matrix

	Height	Circumference	Lean
Height	0.502	0.403	-0.117
Circumference	0.403	0.379	-0.098
Lean	-0.117	-0.098	0.103

C. Standard deviation of prediction

The response surface of the prediction standard deviation is the same basic shape as the prediction variance surface (Figures IV-1 and IV-3). There are some differences in gradients between the surfaces. The extreme gradient between the economic value coordinates, circumference -1.0 lean -10.0,and coordinates, circumference 10.0 lean -10.0, in the prediction variance surface is reduced in the standard deviation surface (Figures IV-1 and IV-3.). The change in surface gradients can be explained by the properties of the square roots of numbers between 0 and 100.

D. Standard deviation of genetic worth

The response surface for the standard deviation of genetic worth (Figure IV-4) has the same basic shape as the response surface for the variance of genetic worth (Figure IV-2). As in the prediction variance and standard deviation

Figure IV-3. Response surface of prediction standard deviation to changes in economic values

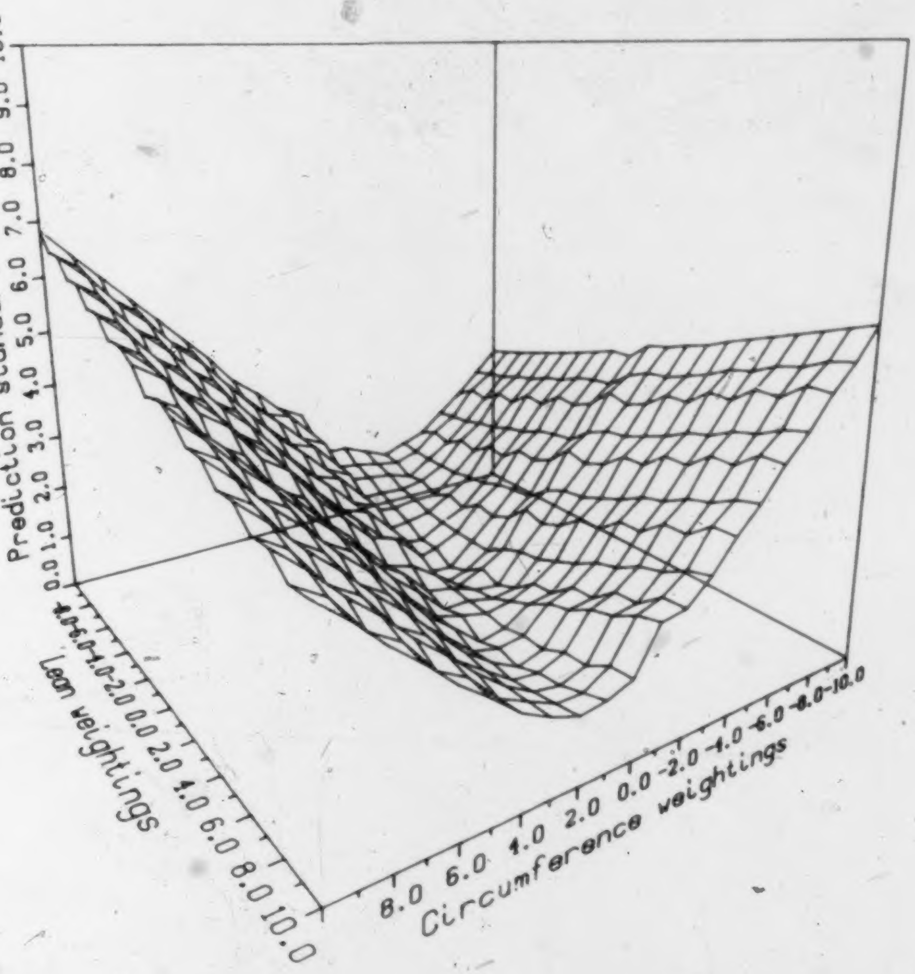
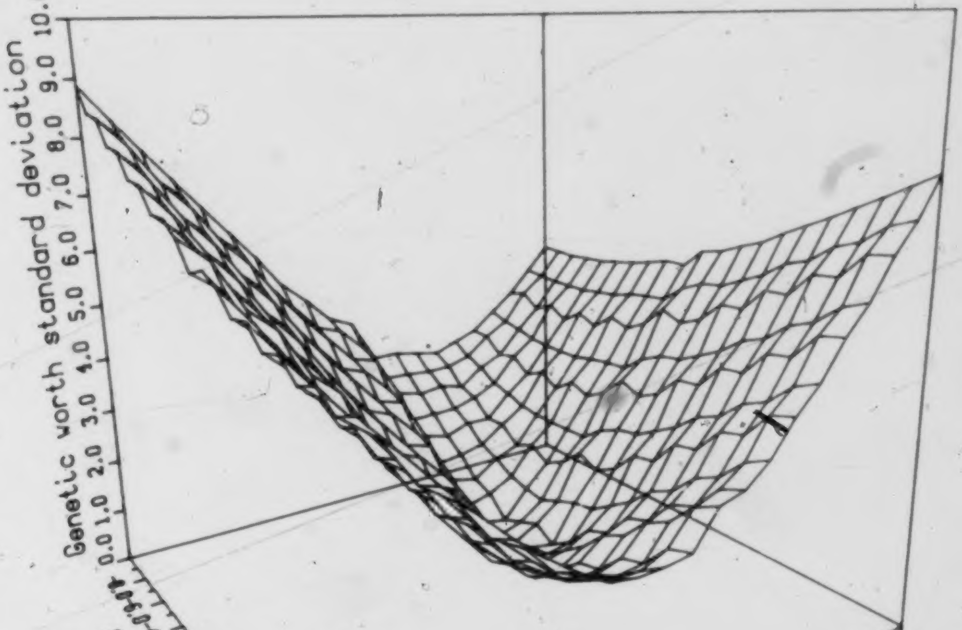


Figure IV-4. Response surface of genetic worth standard deviation to changes in economic values



comparison the gradients are altered.

The most significant difference when the genetic worth standard deviation response surface (Figure IV-4) was compared to the prediction standard deviation response surface (Figure IV-3) occurred in the gradient between economic value coordinates, circumference -2.0 lean 1.0, and coordinates, circumference 2.0 lean 10.0. The genetic worth standard deviation surface rises whereas the prediction standard deviation surface is flat (Figures IV-3 and IV-4). The differences between the surfaces can be explained by the differences between the $\mathbf{C'V^{-1}C}$ and $\mathbf{G}$ matrices. The $\mathbf{C'V^{-1}C}$ components almost cancel each other out in the $\mathbf{a'C'V^{-1}Ca}$ product when the a values have the same values as the coordinates above to give $\mathbf{aC'V^{-1}Ca}$ values between 0.347 and 1.26. Given the scale of the prediction standard deviation the surface jointing the two coordinates appears to be almost flat. The components of the $\mathbf{G}$ matrix almost cancel each other out in the $\mathbf{a'Ga}$ product at economic value coordinates, circumference -2.0 lean 1.0, to give the value of 0.667. The slightly different configuration of the G matrix and the greater value of the components combine to give a much higher variance value (7.67) at economic value coordinates circumference 2.0, lean 10.0. The square root of the genetic worth variance gives the value of the genetic standard deviation at both points on the surface (0.816 and 2.772). Given that the scale of the genetic worth standard deviation is from 0.0 to 10.0 the surface appears to rise

Figure IV-5. Response surface of prediction error variance to changes in economic values

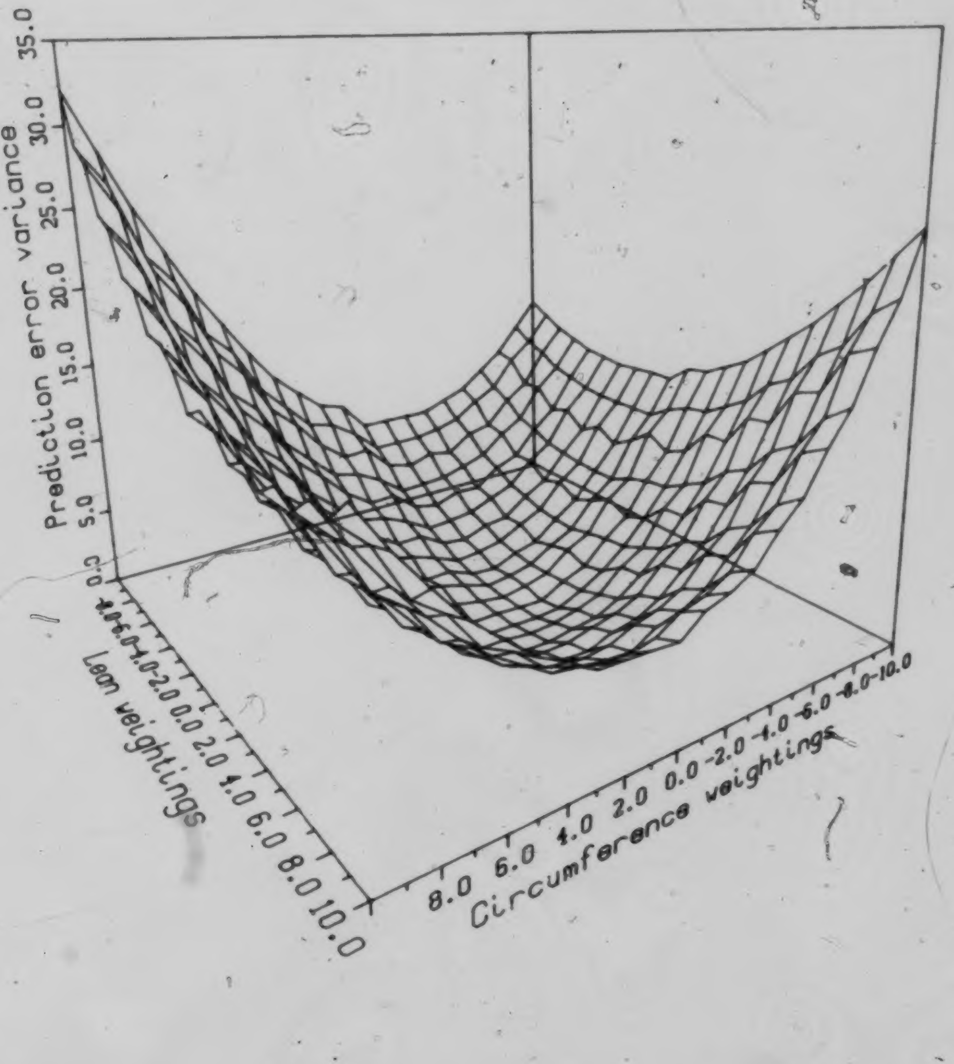
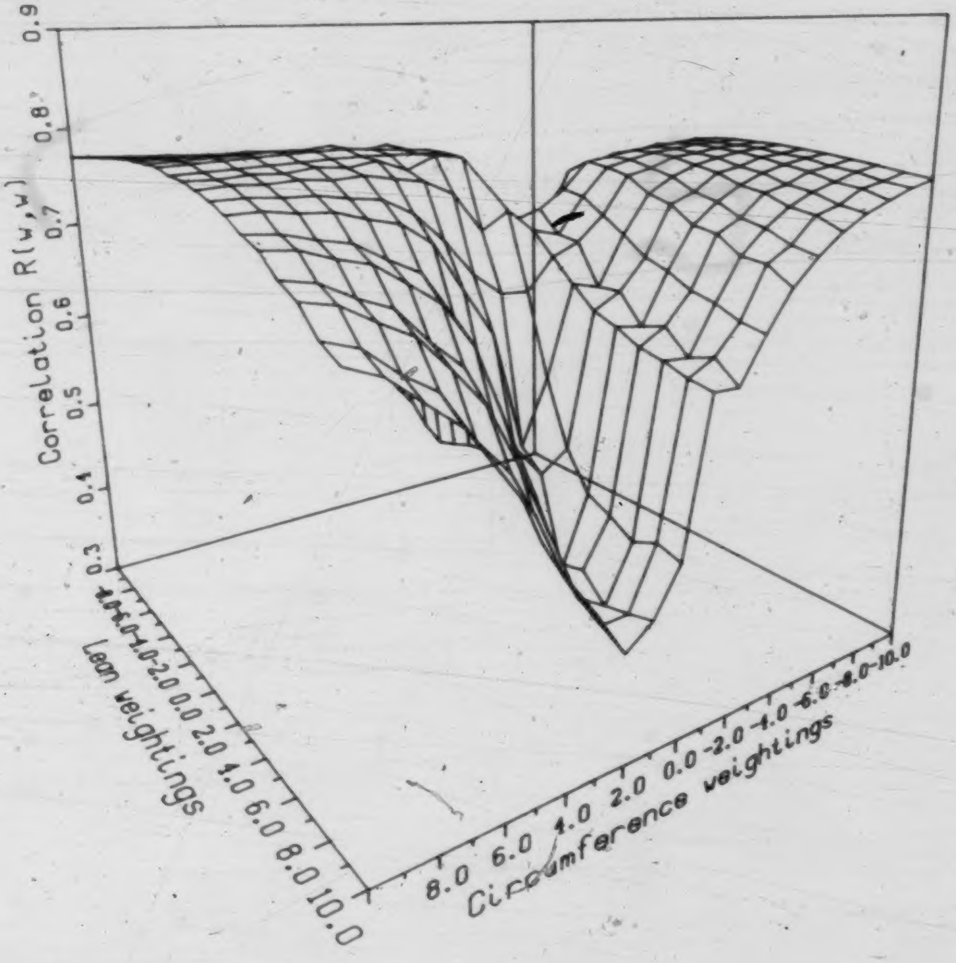


Figure IV-6. Response surface of correlation R(w,w) to changes in economic values



going from standard deviation values 0.816 to 2.772 (Figure IV-4).

E. Prediction error variance

The response surface for the prediction error variance shows that the components of the $C'V^{-1}C$ and G have the same proportions. The difference between the $C'V^{-1}C$ and G matrix components is exagerated as the effect of economic value increases, the highest prediction error variance values occur at the extreme economic values (Figure IV-5).

F. Efficiency

The shape of the response surface for $R(w,\hat{w})$ is alarming because of the areas where the correlation drops sharply and the importance breeders place on the value of the correlation (Figure IV-6). The correlation drops sharply at extreme lean economic values and intermediate circumference economic values.

The drop in the correlation can be explained by the differences between the structure of the G and $C'V^{-1}C$ matrices which caused the differences in response surfaces observed when comparing the differences between the prediction standard deviation and the genetic worth standard deviation response surfaces. The structure of the G matrix produces a relatively large genetic standard deviation value at the economic value coordinates, circumference +2.0 lean -10.0. The difference between the standard deviations at this point produces a low correlation value. There are greater differences between the prediction standard deviation and the genetic worth standard deviation at other economic value coordinates. The important factor is the low prediction variance coupled with the relatively large difference between the prediction standard deviation and genetic worth standard deviation.

The highest correlation occurs at economic value coordinates, circumference 10.0 lean -10.0 (Figure IV-6). This is also the point of highest prediction variance (Figure IV-1). The difference between the prediction variance and the genetic worth is substantial. The difference between the prediction standard deviation and the genetic worth standard deviation is much smaller because of the properties of square roots. In this case the prediction variance is 46.7 and the genetic worth variance is 78.7. The difference between the two values is reduced by taking the square roots. The prediction variance increases as economic values increase. The genetic worth variance increases in the same proportion unless there are differences between the G and $C'V^{-1}C$ matrices. The higher the prediction variance the greater would be the reduction in the difference going from variance to standard deviation. It can be concluded that the economic value coordinates which give the highest prediction variance will give the highest correlation ($R(\hat{w},w)$).

Without a knowledge of the G and $C'V^{-1}C$ matrices it is not possible to predict in advance where the area of low correlation ($R(w,\hat{w})$) will be on the response surface. The data analysed in this contribution shows that if certain economic values are choosen the index may be inefficient. Areas of potential inefficiency are characterised by:

1. Low prediction variance
2. Low genetic worth variance values but relatively high when compared with prediction variance.

The above results are for standardized data. The scale transformation from standardization gives a range of potential prediction variance values between 1 and 100. It may be potentially interesting to transform the data so the lowest prediction variance has a greater numerical value. The properties of square roots would reduce the difference between prediction variance and genetic worth variance to give higher $R(\hat{w},w)$ values.

G. Choice of economic values

Many breeders especially those associated with improving organisms which have shorter production cycles than trees may consider that economic values can be determined with some accuracy. However in forestry there is considerable uncertainty over future economic values. Additional information on index efficiency and prediction variance will become more important when the precision of the economic value prediction is low.

The properties examined in this contribution provide several potential criteria on which the choice of economic values can be based. The most logical parameter to maximize is the correlation R($\hat{w}$,w) because of its relationship to efficiency. There are other properties such as prediction variance and genetic worth variance which are components in the R($\hat{w}$,w) calculation and can be related to optimum efficiency values.

If the breeder is certain of the economic values he or she will just have to accept the r(w,$\hat{w}$) value. However if there is a range of possible economic values because of a lack of precision in the prediction of future economic values the breeder may consider a slight change in economic values to improve r(w,$\hat{w}$). To obtain an efficient index the breeder should select economic values which give high prediction variance. The high prediction variance would avoid potentially inefficient indices.

IV. Conclusion

The index properties were sensitive to potential changes in the economic values. Sensitivity of the properties to changes in economic values can be explained by the structure of the $C'V^{-1}C$ and G matrices. These depend on the traits selected and the breeding population. It is important for breeders to look at the response surfaces and explain them in terms of breeding population parameters. Inefficient indices have low prediction variance and relatively high genetic worth variance. Efficient indices have high prediction variance. The breeder should be aware of potential changes in the $r(\hat{w},w)$ correlation.

V. Literature cited

Baradat, P. 1979 Selection combinee multicaractere chez le pin maritime. Divers modeles d'index de selection utilizes. 104 eme Congres de Societes Savantes, Bordeaux. Fasc. 2. pp. 299-314.

Bulmer, M.G. 1985 The mathematical theory of quantitative genetics. Clarendon Press, Oxford. 254 pages.

Harris, D.L. 1964 Expected and predicted progress from index selection involving estimates of population parameters. Biometrics 20:46-72.

Hayes, J.F. and W.G. Hill. 1980 A reparameterization of a genetic selection index to locate its sampling properties. Biometrics 36:237-248.

Henderson, C.R. 1973 Sire evaluation and genetic trends. *In* Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. J.L. Lush Blacksburg, Virginia. pp. 10-41.

Kendall, M.G. and Stuart, A. 1958 The advanced theory of statistics. Volume I. Distribution theory. Griffin and Company Limited, London. 433 pages.

Rao, C.R. 1965 Linear statistical inference and its applications. Wiley and sons inc., New York. 522 pages.

Smith, C. 1983 Effects of changes in economic weights on the efficiency of index selection. Journal of Animal Science. 56:1057-1064.

Vandepitte, W.M. and L.N. Hazel. 1977 The effect of errors in economic weights on the accuracy of selection indexes. Ann. Genet. Sel. Anim., 9:87-104.

CHAPTER FIVE

SELECTION STRATEGIES INCORPORATING RISK AND UNCERTAINTY OVER ECONOMIC VALUES IN TREE BREEDING

I. Introduction

In multiple trait selection the tree breeder must choose the economic weights at the time of selection. In theory these economic weights should reflect economic values at the time when the improved crop is to be harvested. There is uncertainty over determining the economic value at the time when the improved crop will be harvested. The breeder takes a risk in choosing a set of economic weights at the time of selection. There are decision making techniques to minimize the effects of the uncertainty and reduce risk from the choice of economic weightings (Chernoff and Mosses 1959).

It is the objective of this contribution to apply the appropriate decision making techniques to produce selection strategies which incorporate risk and uncertainty in future economic values. These strategies will be applied to the selection index pioneered by Baradat (1976).

The breeding population for which the selection strategy was developed is a population of maritime pine (*Pinus pinaster* Ait.) produced by a hierarchical mating design. It was evident that the male and female parents did not have the same genetic variance. Selection index formulae will be derived for populations produced from male parents which are considered independent from the female parents. The traits under selection were the internal wood quality traits sampled from increment cores.

II. Materials and methods

A. Biological material

Increment cores were taken from one tree in each plot of a 10-year-old maritime pine controlled pollination progeny trial. The mating design was hierarchical with each one of thirty male parents mated to nine or fewer different females to produce 236 full-sib families. The families were laid out in 110 incomplete blocks. Maritime pine grown in the region where the data were collected, lean because of the stress of the prevailing wind. In theory the stem should be one half compression wood and one half normal wood (Ohta *et al.* 1985). Increment cores were taken in the direction of maximum lean at breast height. The 3028 increment cores were cut at the pith and the specific gravity of each piece determined by Smith's maximum moisture content method (Smith 1954). Three variables were then determined for each sampled tree:

1. Density of the increment core portion with the minimum density.
2. Density of the increment core portion with maximum density,
3. Mean density of the increment core.

These variables will be refered to as minimum, maximum and mean density. The maximum density was considered to be the density of the compression wood portion of the increment core. The minimum density was considered to be the density of the portion of the increment core which is free of compression wood. Logically the breeder would want to increase the minimum and the mean density while reducing the maximum density. Therefore when considering multiple trait selection, reductions in maximum density and increases in minimum and mean densities will be considered desirable.

B. Statistical analysis

There are a number of symbols required in the text. The first time a symbol occurs it is defined. A list of symbols is also given in Table V-1.

The basic input values for selection index calculations are random phenotypic effects and variance components estimated in the analysis of variance. Fixed effects were removed from each record. The linear model of the ajusted record is:

$$Y_{ijk} = \mu + M_i + F_{j(i)} + e_{k(ij)} \quad \ldots(V-1)$$

where Y_{ijk} is the adjusted record value,

M_i is the effect of the male family i,

$F_{j(i)}$ is the effect of the female family j nested within paternal family i, and

$e_{k(ji)}$ is the residual effect of the k individual nested within the j^{th} maternal family.

Table V-1. List of symbols

Symbol	Explanation
Y_{ijk}	record value of the k^{th} individual in the j_{th} female family nested within the i^{th} male family
μ	overall mean
M_i	effect of the i^{th} male family
$F_{j(i)}$	effect of the j^{th} female family nested within the i^{th} male family
$e_{k(ij)}$	residual effect of the k^{th} individual nested within the j^{th} female family
$\hat{p}i$	estimated phenotypic value of the i^{th} male family
$\hat{p}j$	estimated phenotypic value of the j^{th} female family nested within the i^{th} male family
$\hat{p}k$	estimated within family phenotypic value of the k^{th} individual nested within the j^{th} female family
$\hat{v}$	vector of predicted additive genetic values
C	matrix of covariance between phenotypic predictors and the additive genetic values
$\underset{\sim}{p}$	vector of phenotypic predictors
$\hat{Y}_{...}$	average record value
$\hat{Y}_{i..}$	average record value within the i^{th} male family
$\hat{Y}_{ij.}$	average record value within the j^{th} female family nested within the i_{th} male family
$n_{i..}$	number of individuals in the i^{th} male family
$n_{ij.}$	number of individuals in the j^{th} female family nested in the i^{th} male family
nf_i	number of female families nested in the i^{th} male family
AF_{ijk}	additive genetic value from male parent
i	selection intensity
a_S	economic value vector at the time of selection
a_H	economic value vector at the time of harvest
W^*	genetic worth at the time of harvest
$\hat{W}^*$	predicted genetic worth at the time of harvest
e	residual
W	genetic worth at the time of selection
$\hat{W}$	predicted genetic worth at the time of selection

The data was analysed using the UANOVA program developed at the University of Alberta by Dr T. Taerum (Appendix V). The phenotypic effects and variance components output from UANOVA provided the basic data for the construction of the selection indices.

Table V-2. Variance-covariance of male family

	Minimum density	Mean density	Maximum density
Minimum density	31.03	29.77	27.65
Mean density	29.77	28.45	26.28
Maximum density	27.65	26.28	24.10

Table V-3. Variance-covariance of female family

	Minimum density	Mean density	Maximum density
Minimum density	17.47	17.17	16.31
Mean density	17.17	16.59	15.39
Maximum density	16.31	15.39	13.83

The male and female family variance components were different (Tables V-2 and V-3). In theory these variance-covariance matrices should be similar. The maternal variance should be slightly greater because of the addition of dominance variance. It was concluded that the male parents had different genetic variance from the female parents.

C. Modification of prediction formula

Baradat's selection index uses phenotypic predictors to predict genetic values for multiple traits. The predicted genetic values are then economically weighted to give the predicted genetic worth of each individual.

In the heirarchical mating design genetic values are predicted from three phenotypic predictors:

1. The male family phenotypic predictor.
2. The female family phenotypic predictor.
3. The within-family phenotypic predictor.

Each phenotypic predictor can be estimated from the record values:

$$\hat{p}i = \bar{Y}_{i..} - \bar{Y}_{...} \quad ...(V-2),$$

$$\hat{p}j = \bar{Y}_{ij.} - \bar{Y}_{i..} \quad ...(V-3),$$

$$\hat{p}k = Y_{ijk} - \bar{Y}_{ij.} \quad ...(V-4),$$

where $\hat{p}i$ is the estimated phenotypic value of the i^{th} male family,

$\hat{p}j$ is the estimated phenotypic value of the j^{th} female family nested within the i^{th} male family,

$\hat{p}k$ is the estimated within family phenotypic value of the k^{th} individual nested within the j^{th} female family,

$\bar{Y}_{...}$ is the average record value,

$\bar{Y}_{i..}$ is the average record within the i^{th} male family,

$\bar{Y}_{ij.}$ is the average record value within the j^{th} female family nested within the i^{th} male family,

Y_{ijk} is the record value of the k^{th} individual nested within the j^{th} female family.

The formula for the prediction of genetic values is:

$$\hat{v} = C'V^{-1}p \quad ...(V\text{-}6),$$

where $\hat{v}$ is the predicted additive genetic value vector,

C is the covariance between the phenotypic predictors and the genetic values,

V^{-1} is the inverse of the variance-covariance matrix of the phenotypic predictors, and

p is a vector of phenotypic predictors.

i. Calculation of covariances between genetic values and phenotypic predictors

The covariances between the phenotypic predictors and the additive genetic values (C) requires a modification of the basic formulae derived for the hierarchical mating design (Baradat 1982). The additive genetic value of the individual due to the male parent contribution is considered to be independent of the additive genetic value of the individual due to the female parent contribution. The covariance between the individual's additive genetic value due to its female parent and the phenotypic predictors can be calculated by expressing the expectation value of the phenotypic predictors in terms of the individual phenotypic values which contribute to the phenotypic predictor. The covariance between the individual phenotypic values and the maternal additive genetic values can be expressed in terms of the maternal additive variance. The same principle can be used to determine the covariance between the individual's additive genetic value due to its female parent and the phenotypic predictors.

To express the covariance in terms of the additive variance the covariance between the individual's phenotypic value and the additive value to be predicted is based on certain rules derived from basic genetic theory:

1. The covariance between the paternal additive value of an individual and the phenotypic value of that individual is 0.5 of the paternal additive variance.
2. The covariance between the maternal additive value of an individual and the phenotypic value of that individual is 0.5 of the maternal additive variance.
3. The covariance between the paternal additive value of an individual with the phenotypic value of any individual belonging to the same female family is 0.25 of the paternal additive variance.
4. The covariance between the paternal additive value of an individual and the individual phenotypic value of any individuals belonging to the same male family is 0.25 of the paternal additive genetic variance.
5. The covariance between the maternal additive value of an individual and the individual phenotypic value of any individual in the same female family is 0.25 of the maternal additive genetic variance.

6. The covariance between the maternal additive value of an individual and the individual phenotypic value of any of the individuals not belonging to the same maternal family is zero.

Consider the covariance between the three phenotypic predictors and the maternal additive genetic value of individual ijk. The expectation value of the covariance between the individual's maternal additive value and its estimated male family phenotypic value is:

$$\mathrm{Cov}(AF_{ijk}, \hat{Y}_{i..} - \hat{Y}_{...}) \quad ...(V-7),$$

where AF_{ijk} is the maternal additive genetic value of individual ijk.

The $\hat{Y}_{i..} - \hat{Y}_{...}$ is the average of $n_{i..}$ individual phenotypic values. One of these values is the individual phenotypic value of the ijkth individual. The covariance between this individual and the maternal additive value to be predicted is $0.5\sigma_{AF}^2$. There are $n_{ij.} - 1$ individual phenotypic values which belong to the same half sib family. The covariance between each of these individuals and the additive value to be predicted is $0.25\sigma_{AF}^2$.

The expectation value of the covariance between the phenotypic predictor and the maternal additive value can now be rewritten as:

$$\frac{.5 + (n_{ij.} - 1).25}{n_{i..}} \sigma_{AF}^2 \quad ...(V-8),$$

where $n_{i..}$ is the number of individuals in the i^{th} male family,

$n_{ij.}$ is the number of individuals in the j^{th} female family, and

σ_{AF}^2 is the maternal genetic variance.

To express the expectation value of the covariance between the female family phenotypic value and the maternal additive genetic value of an individual in terms of the maternal additive genetic variance the female family phenotypic value is rewritten as:

$$\hat{Y}_{ij.} - \hat{Y}_{i..} = (\hat{Y}_{ij.} - \hat{Y}_{...}) - (\hat{Y}_{i..} - \hat{Y}_{...}) \text{ ...(V-9).}$$

The expected covariance between the maternal additive value and the individual phenotypic values above is:

$$\mathrm{Cov}(AF_{ijk}, \hat{Y}_{ij.} - \hat{Y}_{...}) - \mathrm{Cov}(AF_{ijk}, \hat{Y}_{i..} - \hat{Y}_{...}) \text{ ...(V-10).}$$

Applying the logic and the genetic relationships derived above these covariances can be rewritten in terms of the maternal additive variance:

$$\left[\frac{.5 + (n_{ij.} - 1).25}{n_{ij.}} - \frac{.5 + (n_{ij.} - 1).25}{n_{i..}} \right] \sigma_{AF}^2 \text{ .(V-11).}$$

The expected covariance between the within-family phenotypic value and the maternal additive genetic variance is:

$$\left[.5 - \frac{(.5 + (n_{ij.} - 1).25)}{n_{ij.}} \right] \sigma_{AF}^2 \text{ ...(V-12).}$$

The covariances between the paternal additive values and the phenotypic predictors are determined in the same way as the covariances with maternal additive values. The expected covariance between the paternal additive value of the individual ijk and its male family phenotypic value is:

$$\frac{.5 + (n_{i..} - 1).25}{n_{i..}} \sigma_{AM}^{2} \quad ...(V-13),$$

where σ_{AM}^2 is the paternal additive genetic variance.

The expected covariance between the paternal additive value of the ijkth individual and the female family phenotypic value is:

$$\left[\frac{.5 + (n_{ij.} - 1).25}{n_{ij.}} - \frac{.5 + (n_{i..} - 1).25}{n_{i..}}\right] \sigma_{AM}^{2} \quad (V-14).$$

The expected covariance between the paternal additive value of the ijkth individual and the within-family phenotypic value is:

$$\left[.5 - \frac{(.5 + (n_{ij.} - 1).25)}{n_{ij.}}\right] \sigma_{AM}^{2} \quad ...(V-15).$$

In multiple trait selection the additive genetic variance in the above equations is replaced by the genetic covariance for the non diagonal elements in the sub-matrices which make up the C matrix. The V and C matrices change as the individual's male and female family changes.

ii. Calculation of elements of V matrix

The three phenotypic predictors are considered independent. The V^{-1} matrix is therefore diagonal or block diagonal in multiple trait selection. The individual elements of V can be calculated if it is assumed that the variance of male families, female families and individuals within families are known. The expectation of the variance of the phenotypic values can be expressed in terms of the variance components in exactly the same way as the expectation value of the sum of squares is determined to estimate variance components in the analysis of variance (Searle 1972). The variance of the phenotypic predictors can thus be expressed as:

$$\sigma_{\beta_i}^{2} = \frac{\left[n_{i..}^{2}\sigma_{M}^{2} + \sum_{j=1}^{nf} n_{ij.}^{2}\sigma_{F(M)}^{2} \right]}{n_{i..}^{2}} + \frac{1}{n_{i..}}\sigma_{e}^{2} \quad \text{...(V-16)},$$

$$\sigma_{\beta_j}^{2} = \frac{\left[n_{i..}^{2} - \sum_{j=1}^{nf} n_{ij.}^{2} \right]}{n_{i..}^{2}}\sigma_{F(M)}^{2} + \left[\frac{1}{n_{ij.}} - \frac{1}{n_{i..}} \right] \sigma_{e}^{2} \quad \text{.(V-17)},$$

$$\sigma_{\beta_k}^{2} = \frac{n_{ij.} - 1}{n_{ij.}}\sigma_{e}^{2} \quad \text{...(V-18)},$$

where nf is the number of female families in the i^{th} male family.

All the variance component subscripts in equations V-16, V-17 and V-18 relate to the linear models in equation V-1, V-2, V-3 and V-4. In multiple trait selection the off diagonal elements of the diagonal blocks of V^{-1} are calculated from the same formula. Covariance components are substituted for variance components in equations V-16, V-17 and V-18.

III. Incorporation of decision making theory into selection of economic values

A. Decision theory

Decision theory assumes that the decision maker has a number of strategies. These strategies have several potential outcomes depending on the circumstances or state of nature. In the context of tree breeding the strategy represents the economic weightings at the time of selection. The state of nature represents the economic weightings at the time the improved crop is harvested. The outcomes are evaluated in terms of gain in genetic worth. For example if a breeder chooses strategy n, and state of nature p occurred at the point of harvest, the expected gain in genetic worth would be the value in cell np of Table V-4.

Table V-4. Representation of outcome values for n strategies and p states of nature.

		Strategy	
		1...	...n
	1	1,1	1,n
State of Nature			
	p	p,1	p,n

There are a number of decision making techniques which have been developed for the case when there is risk and uncertainty concerning the outcome of a particular strategy. This contribution will deal with the application of minmax, Bayes, minmax regret and Bayes regret techniques to the choice of optimum economic weightings in genetic selection.

Minmax theory is pessimistic in nature and is based on the sure-thing principle (Savage 1954). The sure-thing principle has been developed into the minmax rule for selection among competing strategies. This rule is best demonstrated by reference to Table V-4. Strategies 1 to n are evaluated on the minimum gain obtainable in the worst possible outcome for each strategy. For example if state of nature 'p' gave the lowest outcome for the first strategy then the value for the first strategy would be the outcome in the position 1,p in Table V-4. The optimum strategy would be the strategy which had the highest minimum value and would be referred to as the minmax strategy.

If the a priori probabilities of the states of nature can be determined then the decision maker can calculate Bayes strategy. The strategy is given a value which is a linear combination of the values of each potential outcome multiplied by the a priori probability. The optimum strategy is the strategy with the highest value and is called Bayes strategy.

Both minmax and Bayes strategies can be calculated on the basis of regret. In this case the values of the outcomes in Table V-4 are assessed in terms of regret. Regret is expressed as the difference between the outcome of the strategy to be evaluated and the strategy which gives the maximum outcome value for the state of nature. The objective is to choose the strategy which minimizes regret. For the minmax philosophy the outcome which gives the highest regret is assigned to the strategy. The strategy with the lowest value is selected as the minmax regret strategy. The strategy values for the Bayes regret technique are calculated in the same way as for the Bayes technique except regret values are assigned to the outcomes for the states of nature. The strategy value with the lowest regret is the Bayes regret strategy.

Minmax strategies have been proposed for decision making in tree breeding (Namkoong 1981). There is a problem because most techniques assume that the function giving the outcome value (gain) is linear. With respect to changes in economic weightings gain in genetic worth is not a linear function. This contribution treats economic weightings as discret units in order to avoid the problems of the non-linearity of the gain function.

B. Calculation of strategy values

The calculation of gain in genetic worth for each possible outcome for a given strategy requires a modification of the gain formula. Gain in a variable, say x, which is assumed to be jointly normally distributed with another variable y, from selection based on y values is given by the formula:

$$\text{Gain}=i\frac{\sigma_{x,y}}{\sigma_y} \quad \ldots(\text{V-19}),$$

where i is the selection intensity,

$\sigma_{x,y}$ is the covariance between x and y, and

σ_y is the standard deviation of y.

When selection indices which predict genetic worth are used to rank individuals the gain in genetic worth is given by the formula:

$$i\,\frac{\sigma_{w\hat{w}}}{\sigma_{\hat{w}}} \quad \ldots(\text{V-20}),$$

where $\sigma_{w\hat{w}}$ is the covariance between the predicted genetic worth based on the economic values at the time of selection and the genetic worth based on the economic value at the time when the improved crop is harvested, and

$\sigma_{\hat{w}}^2$ is the standard deviation of the predicted genetic worth based on the economic values at the time of selection.

To calculate the gains in genetic worth for Table V-4 it is necessary to adapt the formula for the gain in genetic worth. For a given strategy $\sigma_{\hat{w}}^2$ of the gain formula is easily determined. The variance of the predicted genetic worth based on the economic values at the time of selection is:

$$a_s'C'V^{-1}Ca_s \quad ...(V\text{-}21),$$

where a_s is the economic weight vector at the time of selection.

Calculation of $\sigma_{w\hat{w}}$ is more complex. The covariance between the predicted genetic worth based on the economic values at the time of selection and the genetic worth at the time when the improved crop is harvested requires some mathematics but the basic logic has already been expounded in statistical genetics texts (Bulmer 1985). The genetic worth of an individual based on the economic values at the time of harvest is:

$$w^* = \hat{w}^* + e \quad ...(V\text{-}22),$$

where w^* is the genetic worth at the time of harvest,

$\hat{w}^*$ is the predicted genetic worth based on the economic values at the time of harvest, and e is the residual.

The genetic worth of an individual based on the economic values estimated at the time of selection is:

$$w = \hat{w} + e \quad \ldots\text{(V-23)}$$

where w is the genetic worth based on the economic values estimated at the time of selection,

$\hat{w}$ is the predicted genetic worth based on economic values estimated at the time of selection, and

e is the residual.

The expectation values $E(w)$, $E(w^*)$, $E(\hat{w})$ and $E(\hat{w}^*)$ are by definition all zero. Therefore the covariance between w^* and $\hat{w}$ is given by multiplying equation (V-22) by $\hat{w}$ and taking expectation values to give:

$$\mathrm{Cov}(w^*,\hat{w}) = \mathrm{Cov}(\hat{w}^*,\hat{w}) \quad \ldots\text{(V-24).}$$

The covariance $(\hat{w}^*,\hat{w})$ can be rewritten in terms of the basic variance-covariance matrices of the selection index:

$$a_H'C'V^{-1}Ca_S \quad \ldots\text{(V-25),}$$

where a_H is the vector of economic weightings at the time of harvesting.

The outcome values can be calculated by the formula:

$$\text{Gain} = i\frac{a_H'C'V^{-1}Ca_S}{(a_S'C'V^{-1}Ca_S)^{.5}} \quad \ldots\text{(V-26).}$$

For a given strategy a_s is constant and a_H changes as the state of nature changes. Strategy values were calculated for two sets of economic values: firstly, a range in which both negative and positive econonmic values were included (Table V-5), and secondly, a range in which only positive economic values were included (Table V-6). The range of economic values for the states of nature is the same as the range of economic values for the strategies. The logic is that a rational individual would only choose those economic values which he or she considered possible at the time of harvest.

Table V-5. First set of economic weights (minimum density held constant)

Mean density	-10	-8	-6	-4	-2	1	2	4	6	8	10
Maximum density	-10	-8	-6	-4	-2	1	2	4	6	8	10

Table V-6. Second set of economic weights (minimum density held constant)

Mean density	0	2	4	6	8	10	12	14	16	18	20
Maximum density	0	2	4	6	8	10	12	14	16	18	20

Numerical examples of the calculation of the C and V matrices and the determination of stragety values are given in appendix IX.

IV. Results and discussion

The results are specific to one breeding population, but they are also of interest to all breeders as they can provide information which is relavent to their own programs. All procedures were repeated for selection intensities of 1.4 and 1.96. Strategy values were plotted against economic weightings to give additional information on the sensitivity of departures from optimum strategies. The response surfaces were the same for both selection intensities; the only difference was that the outcome values were numerically greater for the selection intensity of 1.96. The results presented here are for the 1.4 selection intensity, because the response surface of the optimum outcome values for each strategy has more gradual gradients and is more easily interpreted.

Table V-7. Economic weightings of the optimum strategies for first set of economic values

	Minmax	Minmax regret	Bayes	Bayes regret
Mean density	8	-10	2	2
Maximum density	10	-10	1	1

Table V-8. Economic weightings of the optimum strategies for second set of economic values

	Minmax	Minmax regret	Bayes	Bayes regret
Mean density	4	6	1	1
Maximum density	6	8	1	1

In theory the standard and regret strategies should be the same (Blackwell and Girshick 1954). The Bayes strategy is the same as the Bayes regret strategy for both sets of economic weightings (Tables V-7 and V-8). The minimax strategy differs from the minmax regret strategy for the

Figure V-1. Surface of strategy values for minmax technique (first set of economic values)

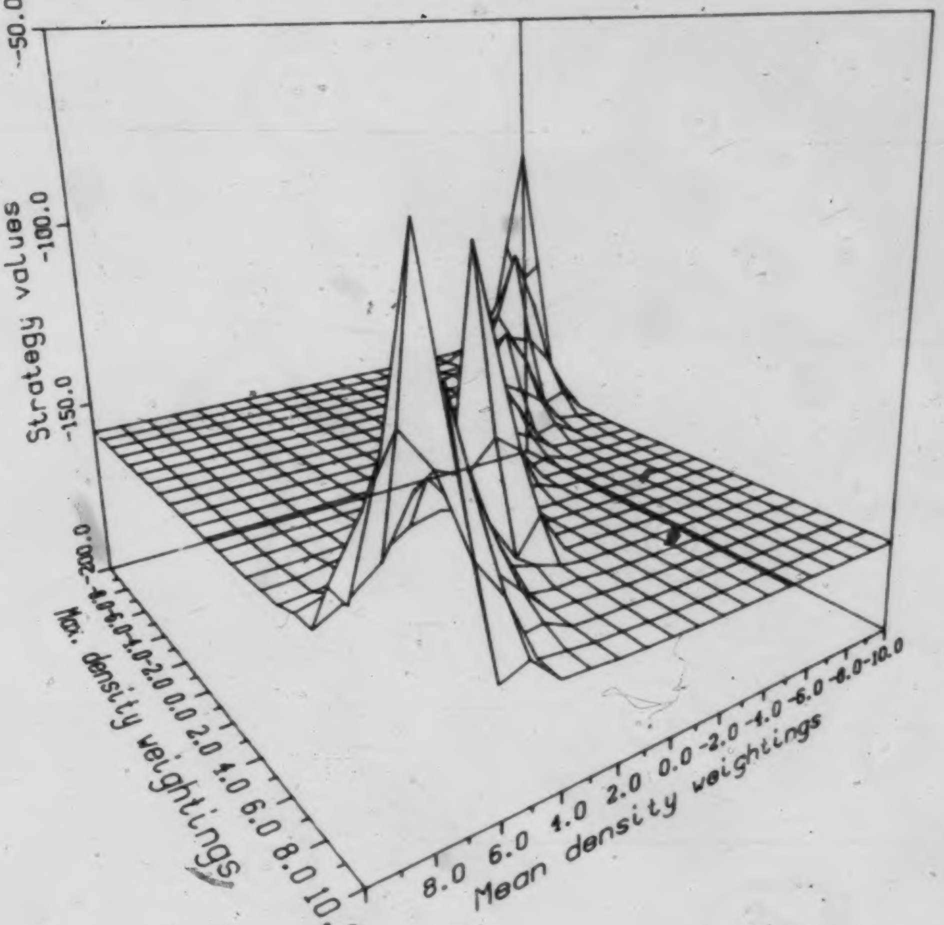
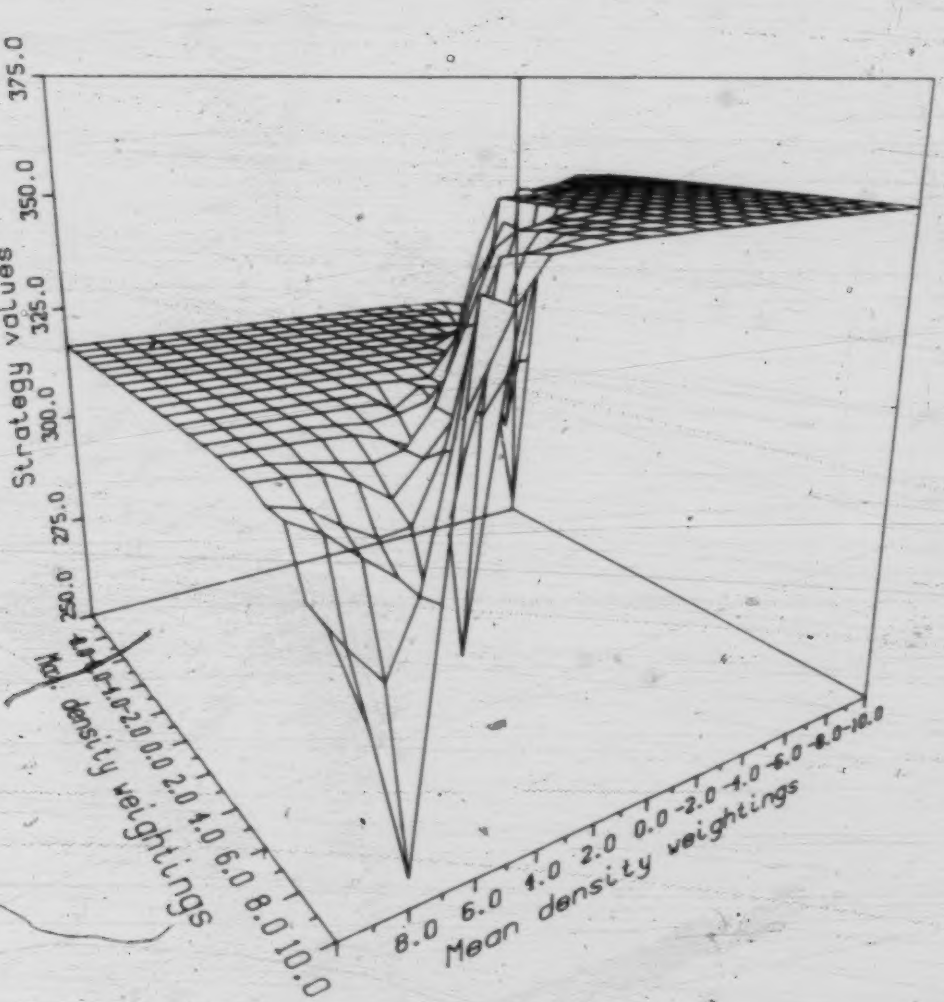


Figure V-2. Surface of strategy values for minmax regret technique (first set of economic values)



first and second set of weightings. The minmax and minmax regret strategies for the first set of weightings are almost exact opposites (Table V-7). In contrast, for the second set of weightings the strategies are similar. An explanation of these departures from theoretical expectations can be seen in the surfaces of the strategy values for the minmax and minmax regret techniques (Figures V-1 and V-2). The surface of the minmax regret strategy values is a mirror image of the minmax strategy values surface. Both surfaces have three peaks: one for positive economic weightings, one for negative economic weightings and one for the transition between positive and negative economic weightings (Figures V-1 and V-2). The minmax regret technique selected the optimum strategy associated with the peak for the negative economic weightings, while the minmax technique selected the optimum strategy associated with the peak for the positive economic weightings. Outcome values for the minmax regret and regret strategies are very similar at the three peaks on the response surface. The change in optimum strategies between minmax and minmax regret strategies is caused by the assumption that the economic values are discrete integers. The economic weightings for the true optimum strategy are not integer values. The integer economic values give strategy outcome values each side of the optimum strategy outcome. The proximity of the minmax and the minmax regret strategies for the second set of economic values confirms

Figure V-3. Surface of strategy values for minmax technique (second set of economic values)

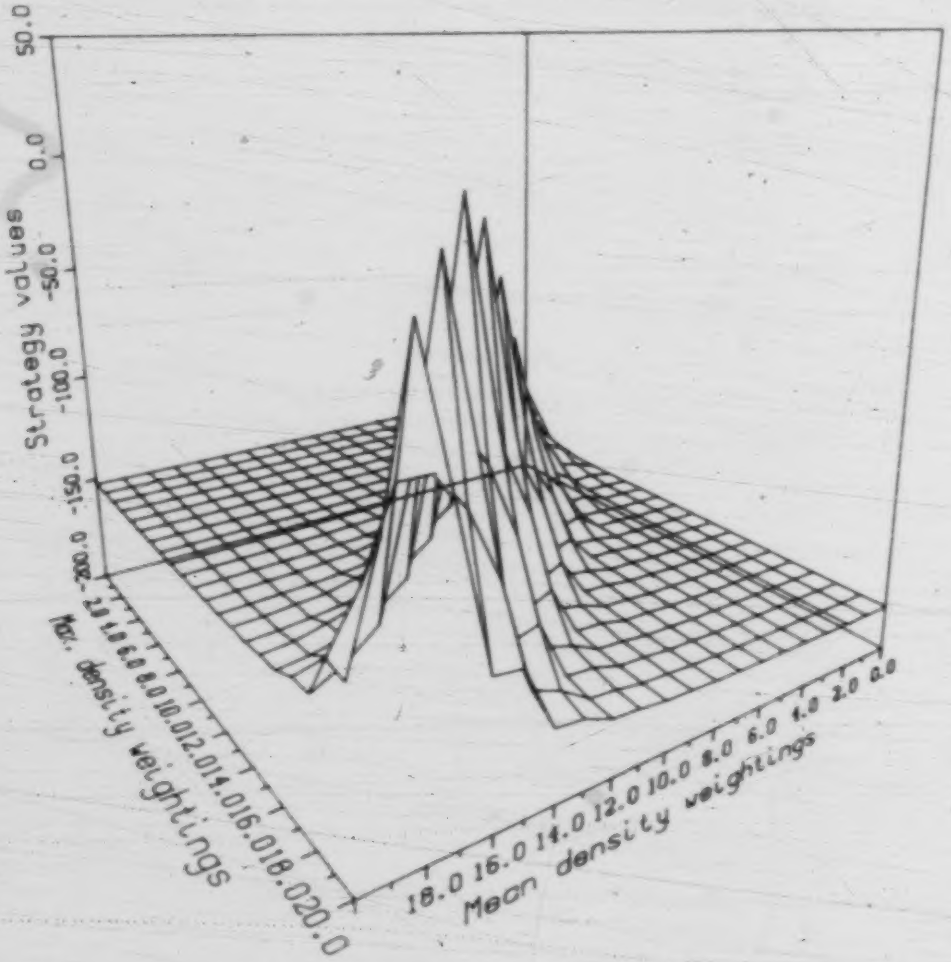
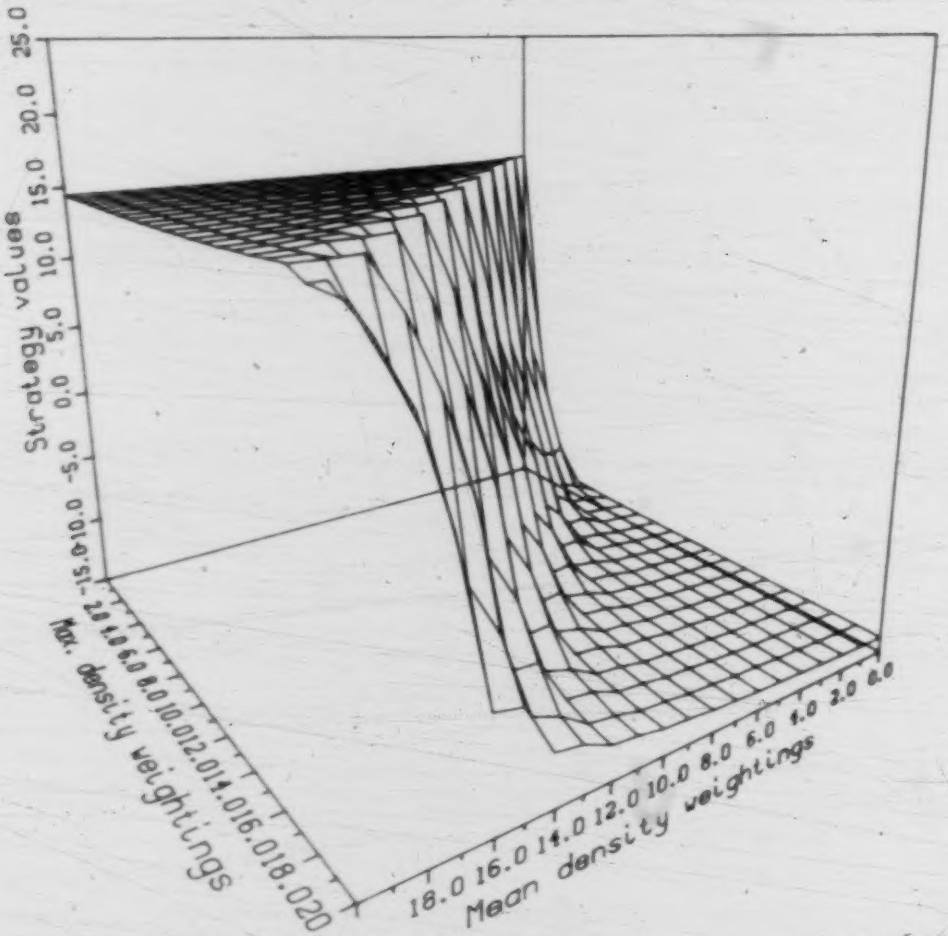


Figure V-4. Surface of strategy values for Bayes technique (second set of economic values)



this explanation. The Bayes and Bayes regret strategies coincide, because the strategy values in the Bayes technique are a weighted average or in this case an average of all the states of nature. This reduces the effects of the assumption that the economic weights are discrete integer values.

A comparison of the surfaces of the strategy values for Bayes and minmax techniques shows that the Bayes optimum value is on a plateau compared with a sharpe peak for the minmax values (Figures V-3 and V-4). The implications are that if the tree breeder accepts the logic of the minmax argument he or she has little flexiblity. On the other hand with the Bayes technique non-optimal strategy values on the same plateau as the Bayes strategy could be choosen without a potential of loss in gains (Figure V-4). It is unlikely that all surfaces of strategy values determined by Bayes technique would be the same shape if different a priori probabilities were choosen.

The optimum strategies for all techniques changed as the range of ecomomic values changed (Tables V-7 and V-8). This would be expected because of the way the outcome values are calculated. It does have significance for the tree breeder. He or she must decide what the potential range of economic values is before applying any of these techniques.

The minmax technique has an advantage because the decision maker only has to decide what are the potential economic weightings. There is no requirement to give probabilities to potential economic weightings. In contrast the Bayes strategy requires a priori probabilities. This requirement may be useful in tree breeding as it will concentrate breeders and decision makers' minds on the relevant questions.

V. Conclusion

The results presented here are for one breeding population, the phenotypic and genetic variances obviously influence the results. The application of decision making techniques to selection decisions is of interest to all breeders. Several conclusions are of interest. Bayes and minmax strategies change as the range of economic values considered changes. The breeder must decide on the range of economic values to be considered. When Bayes strategies are to be determined it is necessary to determine a priori probabilities for the states of nature. Response surfaces showing the values for a range of selection strategies for a given decision making technique give the breeder additional information. The results of the present study indicate that small departures from minimax optimum straegies can result in inefficient selection.

VI. Literature cited

Baradat P. 1976 Use of juvenile-mature relationships and information from relatives in combined multitrait selection. In: Proceedings of the IUFRO joint meeting of genetic working parties on advanced generation breeding. Bordeaux. Institut National de la Recherche Agronomique. pp. 121-138.

Baradat Ph. 1982 Genetique quantitative modeles statistiques et genetiques de base. Institut National de la Recherche Agronomique. 204 pages.

Blackwell, D. and M.A. Girshick. 1954 Theory of games and statistical decisions. Dover Publications. Inc. New York. 345 pages.

Bulmer, M.G. 1985 The mathematical theory of quantitative genetics. Clarendon Press. Oxford. 254 pages

Chernoff, H. and L.E. Moses. 1959 Elementary decision theory. Dover Publications Inc. New York. 364 pages.

Namkoong, G. 1981 Introduction to quantatitive genetics in forestry. Castle House Publications, Tunbridge Wells. 342 pages.

Ohta, S., R. Keller and G. Janin. 1985 Effets de divers modes de fertilisation (N,P,K) sur certaines carateristiques physiques, chimiques mecaniques et proprietes papetieres du pin maritime des Landes (*Pinus pinaster* Ait.) II. Bois de compression et proprietes papetieres. Ann. Sci. For., 42(1)69-96.

Savage, L.J. 1954 The foundations of statistics. Dover Publications Inc. New York. 310 pages.

Searle, R.S. 1972 Linear models. John Wiley and Sons, New York. 532 pages.

Smith, D.M. 1954 Maximum moisture content method for determining specific gravity of small wood samples. United States Forest Service, FPL Report No. 2014.

CHAPTER SIX

PREDICTION METHODS FOR THE FUTURE

I. Introduction

The best prediction method is the one which most accurately predicts the genetic worth of an individual. The method has to accurately predict genetic values and incorporate economic values into the prediction of genetic worth. At the present time tree breeders prefer to use selection index instead of BLUP because it is computationally less demanding than BLUP. When the breeder is only interested in predicting random genetic effects on one site, selection index gives essentially the same results as BLUP (Chapter II). Statistical decision making can be incorporated into the choice of economic values in the selection index calculations (Chapter V).

It is the objective of this contribution to look ahead to the future demands of tree breeding and to look at potential problems with BLUP and selection indices and to propose some solutions to these problems.

II. Future demands of tree breeding

Any predictions of future trends are subject to personal bias. There is obviously a great deal of discussion on the breeding methods to be used, this discussion will be avoided. It will be assumed that breeders will need to predict the performance of genotypes. The predictions of genotype performance will be influenced by five factors.

1. Genotypes will be planted over several different site types.
2. Because of the large number of crosses genotypes will be planted over several years.
3. Populations will be more inbred.
4. There will still be a need to predict gains.
5. There will still be uncertainty over the economic values of the improved crop when it is harvested.

Considering these requirements any future prediction should include site and planting year. In addition it should be possible to incorporate inbreeding and give expected gain values.

III. Mixed model predictions

Mixed models have been used to predict random genetic values in tree breeding (Baradat 1979). Genetic values are predicted after within site (block) fixed effects have been removed. In the future tree breeders will be dealing with more effects, for example, year of planting and site effects. There is some debate over which factors are considered random and which factors are considered fixed. Obviously genetic effects are random. The classification of environmental effects is more debatable.

Consider an example of a breeder who has crossed a large number of trees. Progeny are planted over several years on several different sites. The linear model for the record of an individual is:

$$Y'_{ijklm} = \mu + S_i + A_j + B_k + G_l + e_m \;...(VI-1),$$

where Y'_{ijklm} is the record value of the l^{th} genotype in the k^{th} block on the i^{th} site planted in the j^{th} year,

μ is the mean of all records,

S_i is the effect for the i^{th} site,

A_j is the effect for the j^{th} year,

B_k is the effect for the k^{th} block,

G_l is the effect for the l^{th} genotype, and

e_m is the residual effect.

The site, year and block effects could all be considered random. Burdon (1979) proposed a selection index to predict genetic values when sites and blocks are considered to be random. The theory of the method is sound, but there are two major problems.

1. Future performance values of individuals cannot be predicted.
2. When there is a large number of sites the V matrix would be large and therefore difficult to invert.

A better approach would be to assume that all the environmental effects are fixed. The fixed effects could be estimated by a least squares procedure using the analysis of variance or by the BLUP procedure in which fixed effects are maximum likelihood estimates. Once the individual effects in the linear model have been estimated a predicted performance value can be calculated for each genotype. The predicted performance value is a linear combination of genetic and fixed environmental effects. For example consider a breeder who wants to predict the performance of genotypes on several sites which have been planted over several years. If all the fixed effects are not removed from the record values before the genotypic effects are predicted the predictions will be biased. The year of planting effect cannot be reproduced when the new crop is planted. Considering the fact that the genotypes will be planted over a number of years the year of planting effects in future crops should be very close to zero. Thus the year of planting should be considered in the prediction of genetic values but can be neglected in the prediction of future performance values of genotypes.

Site effects should be included in the prediction of future performance values because the site effects can be reproduced and will affect the average performance values depending on the area of each site type. The effect of block within site is difficult to reproduce although it is considered fixed. Any prediction of performance values must include an average block value for each site. The predicted performance value of a genotype planted on a site is:

$$P\hat{P}V = \hat{\mu} + \hat{S}_i + \hat{B}_{i.} + \hat{G}_i \quad \ldots \text{(VI-2)},$$

where $P\hat{P}V$ is the predicted performance value,

$\hat{\mu}$ is the estimated mean value,

$\hat{S}_i$ is the estimated site effect,

$\hat{B}_{i.}$ is the average of estimated block effects within site i, and

G_1 is the estimated genetic effect of genotype 1.

IV. BLUP versus selection index

Tree breeders have not used BLUP or selection indices extensively. There are a few examples of selection indices in tree breeding (Cotterill 1985). BLUP has not been used because it was developed in animal science and because it requires complex computations.

If either of these methods are to accurately predict the future performance of genotypes, predictions must include unbiased fixed and random effects. In simple experiments rankings on the basis of random genetic values are the same for both BLUP and selection indices (Chapter II). Fixed effects estimates differ, in theory BLUP is more correct.

As tree breeding programs progress breeding populations will become more inbred. Both selection index and BLUP can incorporate inbreeding when calculating random effects (Henderson 1976; Baradat 1982). Inbreeding is incorporated into the estimation of fixed effects in the BLUP calculations (Chapter IV). It is not obvious how inbreeding can be included in the estimation of fixed effects in the analysis of variance calculation.

Flexibility is an important trait of any prediction method. BLUP is extremely flexible because of the flexibility of the basic linear model on which BLUP is based (Henderson 1974). The model allows the breeder to incorporate fixed, random and the interactions between fixed and random effects in predictions (Henderson 1974). Estimation of fixed effects in the analysis of variance and selection index predictions of random effects is a reasonably flexible model. However some development is necessary to incorporate interactions between fixed and random effects into the prediction process.

The numerical analysis problems in calculation of the BLUP estimates are considerable. Many of the problems have been solved (Schaeffer and Kennedy 1986). The solutions to the problems have not yet been adopted in tree breeding because selection index calculations are less demanding. The major effort in selection index calculations is the calculation of phenotypic effects. When numbers are small and data sets are balanced the calculations are relatively simple. As data sets become larger and more imbalanced the computational effort to calculate effects will be greater. For example the effort to calculate the phenotypic effects for 20,000 trees each measured for three traits is considerable. It will be possible to do these calculations but some modifications of existing software will be required. It can be concluded that the numerical analysis problems of selection index calculations will be greater in the future when data sets are larger. The difference between the computational demands of BLUP and selection index predictions will be reduced.

The importance of BLUP in tree breeding programs will increase in the future because of BLUP's flexibility and ability to include inbreeding in mixed model predictions. There are a number of potential problems with BLUP applications in tree breeding programs.

V. Potential problems of implementing BLUP in tree breeding

There are two major problems with BLUP implementation. The first is a numerical analysis problem. The second is that there is no way to calculate expected gain as is done in selection indices.

A. Numerical analysis

There are some important points to be noted in the development of BLUP applications in animal science. The first application of BLUP was for sire evaluation (Henderson 1973). This was single trait evaluation. There were a large number of animals, but the dimension of the random effects part of the coefficients in the BLUP equation is of the order of the number of sires (Henderson 1973). The numerical analysis problems were not overwhelming and the model had immediate application on a national and international scale. Algorithms by Henderson (1976) and Quaas (1976) helped solve some of the numerical analysis problems and allowed the method to be expanded to the prediction of individual animals. Use of iterative solutions to solve for unknowns in the BLUP equations extended capabilities so that equations with more than 100,000 unknowns could be solved (Schaeffer and Kennedy 1986). Development of the methods was largely confined to Cornell University and the University of Guelph. Models developed there had national applications.

In forestry there is a much greater range of types of genetic experiments. Mating designs range from open pollination to complete diallel crosses (Zobel and Talbert 1984). Experimental layouts of field trials cover the complete range from randomized complete block to incomplete latin squares (Montgomery 1984). BLUP models are flexible and in theory models exists which are directly related to tree breeding problems (Henderson 1974, Henderson 1988). These models have not been used in forestry. To understand the BLUP prediction models and computer programs a reasonable level of statistics and numerical analysis is required.

There is a question over the level of numerical analysis in tree breeding. Cotterill (1985) comments on the "lack of special expertise and computing programs required to construct the Smith-Hazel indices". The Smith-Hazel indices refered to by Cotterill were indices for mass selection that are relatively easy to construct. If his assessment is correct problems with BLUP will be enormous. Tree breeders are fully employed establishing trials and taking field measurements. There are two possible solutions:

1. Establishment of a research project to develop BLUP models for use in tree breeding.
2. Evaluation of BLUP models used in animal science with a view to direct transfer to tree breeding.

The solutions to the numerical analysis problems have been determined (Henderson 1976; Schaeffer and Kennedy 1986). The new generation of personal computers should be powerful enough to remove any hardware limitations. It is a question of allocation of resources to implement the solutions.

B. Expected gain calculations and BLUP

BLUP gives unbiased predictions of genetic values. When economic weights have to be included in multiple trait BLUP they are calculated in other experiments and then applied to the genetic values (Henderson and Quaas 1976). BLUP is based on accurate prediction and is not concerned with maximizing gain, however gain is maximized by accurate prediction (Bulmer 1985). Because of the emphasis on accurate prediction no method for calculating gain has been developed. Tree breeders need to be able to calculate gain because:

1. Decision makers who allocate resources in forestry and outside of forestry require some method of calculating potential gains.
2. There is uncertainty over economic values of traits at the time when the new improved crop will be harvested. This uncertainty can be incorporated into decision making when choosing economic values if gain

calculations can give the gains for a set of economic values at the time of selection and for a range of sets of economic values at the time when the improved crop is harvested (Chapter V).

i. Calculation of gain for decision makers

Any gain calculation must take into account the mixed model in the BLUP predictions. Consider an example of selected genotypes planted out in blocks over several sites. If the breeders objective is to maximize gain over several sites the linear model the breeder wants to predict is:

$$\hat{PPV} = \hat{\mu} + \hat{S}_i + \hat{B}_{i.} + \hat{G}_i \quad \ldots(\text{VI-3}).$$

Expected values of a population produced by random mating of the selected population grown on a given site (i) will be:

$$E(\hat{PPV}_T) = \hat{\mu} + \hat{S}_i + \hat{B}_{i.} + E(G_T) \quad \ldots(\text{VI-4}),$$

where $\hat{PPV}_T$ is the predicted performance value of the genotypes in the group selected by truncation selection, and

G_T is the random genetic effects of the genotypes in the selected group.

The general formula for expected gain from truncation selection on the basis of predicted values is:

$$\text{Gain} = i \, \frac{\text{Cov}(x, \hat{x})}{\text{Var}(\hat{x})} \, (\text{Var}(\hat{x})^{.5}) \quad \ldots(\text{VI-5}),$$

where i is the selection intensity;

$Cov(x,\hat{x})$ is the covariance between the prediction and the value to be predicted; and

$Var(\hat{x})$ is the variance of the prediction (Bulmer 1985).

Bulmer showed that:

$$Cov(x,\hat{x}) = Var(\hat{x}) \quad ...(VI-6).$$

Equation (VI-5) can now be rewritten as:

$$Gain = i\ Var(\hat{x})^{.5} \quad ...(VI-7).$$

The variance of $Var(\hat{x})$ could be estimated directly from the BLUP predictions. The expected value of trees gown on a site say site 'i' is given by the linear combination:

$$Site_i + Block\ average_i + Expected\ Gain \quad ...(VI-8).$$

Gains over several sites could be calculation from gains on each site.

ii. Gain formula for selection strategies incorporating uncertainty over economic values

The gain formula for incorporating uncertainty over economic values into selection index is:

$$Gain = i\frac{a_H C'V^{-1}Ca_S}{a_S C'V^{-1}Ca_S}(a_S C'V^{-1}Ca_S) \quad ...(VI-9),$$

where $C'V^{-1}C$ is the variance of the prediction,

a_H is a vector of the economic values at the time of harvest, and

a_S is a vector of the economic values at the time of selection.

The above formula can be easily adapted by replacing the $C'V^{-1}C$ by the variance of the random BLUP effects. The minmax and Bayes strategies can now be determined to give optimum 'a' values on a given site (Chapter V).

VI. Conclusions

The importance of BLUP in tree improvement will increase in the future because BLUP mixed model predictions are flexible and can include inbreeding in fixed effects estimation. There will be problems with the numerical analysis and computing skills required to obtain BLUP predictions. These problems can be solved. Success will depend on the resources allocated by decision makers. Expected gain values to truncation selection for BLUP have not been derived. Formulae used in gain calculations in index selection can be adapted provided the breeder is willing to accept the estimate of the variance of the random BLUP predictions.

VII. Literature cited

Baradat, P. 1982 Genetique quantitative modeles statistiques et genetiques de base. INRA. 204 pages.

Bulmer, M.G. 1985 The mathematical theory of quantative genetics. Clarendon Press, Oxford. 254 pages.

Burdon, R.D. 1979 Generalisation of multi-trait selection indices using information from several sites. New Zealand Journal of Forestry Science 9:145-152.

Cotterill, P.P. 1985 On index selection II. Simple indices which require no genetic parameters or special expertise to contruct. Silvae Genetica 34:64-68.

Henderson, C.R. 1974 General flexibility of linear model techniques for sire evalution. Journal of Dairy Science 57:963-972.

Henderson, C.R. 1976 A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.

Henderson, C.R. and R.L. Quaas. 1976 Multiple trait evaluation using relatives records. Journal of Animal Science 43:188-197.

Henderson, C.R. 1988 Theoretical basis and computational methods for a number of different animal models. *In* Animal Model Workshop, Journal of Dairy Science. Vol. 71 Supplement2:1-16.

Montgomery, D.C. 1984 Design and analysis of experiments. 2nd Edition, John Wiley and sons, New York. 538 pages.

Quaas, R.L. 1976 Computing the diagonal elements and the inverse of a large numerator relationship matrix. Biometrics 32:949-953.

Schaeffer, L.R. and B.W. Kennedy. 1986 Computing strategies for solving mixed model equations. Journal of Dairy Science 69:575-579.

Zobel, B.J. and J.T. Talbert. 1984 Applied forest tree improvement. John Wiley and sons, New York. 505 pages.

APPENDIX I.

DERIVATION OF SELECTION INDEX FORMULA

Consider the problem of multiple trait selection based on an index which combines the phenotypic values from n genetic groups for q traits. The joint distribution of the phenotypic and genetic values are assumed to be multivariate normal with a variance-covariance structure:

$$\begin{matrix} G & C \\ C' & V \end{matrix}$$

where G is a q × q matrix of the variance and covariance
among the q unknown genetic values,
C is a q × (n × q) matrix of the covariances among the
n × q phenotypic predictors and the q unknown
genetic values, and
V is a (n × q) × (n × q) of the variances and
covariances among the n × q phenotypic
predictors.

Breeders are generally concerned with additive genetic values. Henceforth the term genetic value will be defined to mean "additive genetic value".

The genetic worth of an individual with q genetic values can be written as a linear combination:

$$W = a_1v_1 + \ldots + a_qv_q \ldots (AI-1),$$

where W is the genetic worth of the individual,

$a_1 \ldots a_q$ are the relative economic values of the q traits, and

$v_1 \ldots v_q$ are the genetic values of the q traits.

The objective of the breeder is to determine the b values which maximize the genetic worth (W) of the individuals in the selected population. The b values are applied to the phenotypic values to give a linear combination known as the selection index:

$$I = b_{1,1}p_{1,1} + \ldots + b_{n,q}p_{n,q} \quad \ldots(\mathrm{AI}\text{-}2),$$

where I is the index value of the individual,

$b_{1,1} \ldots b_{n,q}$ are the b values for the phenotypic traits, and

$p_{1,1} \ldots p_{n,q}$ are the phenotypic values of the p traits in the n genetic groups.

If the relative economic values of the q traits are known the expected genetic worth of an individual in the breeding population is predicted by a linear function of phenotypic values such that:

$$E(W|I) = \mu_W + B_{W,I}(I - \mu_I) \quad \ldots(\mathrm{AI}\text{-}3),$$

where I is a linear combination of phenotypic values (selection index),

μ_W is the mean genetic worth of the individuals in the population,

$B_{W,I}$ is the regression coefficient of the regression of the linear combination W on the linear combination I, and

μ_I is the average index value for the population.

The mean genetic worth from truncation selection on the basis of selection index values in the base population is:

$$E(\hat{W}_{T.} | \hat{I}_{T.}) = \mu_W + B_{W,I}(\hat{I}_{T.} - \mu_I) \ldots (AI-4),$$

where $\hat{W}_{T.}$ is the mean genetic worth in the selected population,

$\hat{I}_{T.}$ is the mean index value in the selected population, and

μ_W is the mean genetic worth in the base population.

The mean genetic worth in the base population is a constant, therefore the expected gain in genetic worth from truncation selection is:

$$B_{W,I}(\hat{I}_{T.} \ \mu_I) \ldots (AI-5).$$

Standardizing the selection differential gives:

$$B_{W,I} i\sigma_I \ldots (AI-6),$$

where i is the standardized selection differential or selection intensity, and

σ_I is the standard deviation of the selection index.

Expressing the terms in the form of their basic linear components $B_{W,I}$ can be rewritten as:

$$\mathbf{a'Cb(b'Vb)}^{-1} \ldots (AI-7),$$

where **a** is a vector of economic values, and

b is a vector of index coefficients,

and σ_I can be rewritten as:

$$(\mathbf{b'Vb})^{.5} \quad \ldots(\text{AI-8}).$$

Substituting the values for $B_{W,I}$ and σ_I into equation (AI-6) gives:

$$i(\mathbf{a'Cb})(\mathbf{b'Vb})^{-1}(\mathbf{b'Vb})^{.5} \quad \ldots(\text{AI-9}).$$

If the relative economic weights are known the expected gain in genetic worth is maximized by maximizing the correlation between I and W. The proof of this statement is given by multiplying equation (AI-9) by $(\mathbf{a'Ga})^{.5}(\mathbf{a'Ga})^{-.5}$ which gives:

$$(\mathbf{a'Ga})^{.5}(\mathbf{a'Cb})(\mathbf{b'Vb})^{-.5}(\mathbf{a'Ga})^{-.5} \quad \ldots(\text{AI-10}).$$

The last three terms are equivalent to the correlation coefficient between I and W. The first term $(\mathbf{a'Ga})^{.5}$ is constant when the economic values are known. Thus the B values which maximize the value of the correlation between I and W will maximize gain.

The b values which maximize the correlation between I and W are determined by equating the partial derivative of the correlation coefficient to zero giving:

$$\frac{\mathbf{a'C}}{\mathbf{a'Cb}} \; \frac{\mathbf{Vb}}{\mathbf{b'Vb}} = 0 \quad \ldots(\text{AI-11}).$$

All the solutions to the above equations are proportional to the solutions:

$$Vb = C'a \quad ...(AI-12).$$

Premultiplying both sides by V^{-1} gives the classic equation for the b values of the selection index:

$$b = V^{-1}C'a \quad ...(AI-13).$$

APPENDIX II.

EQUIVALENCE OF BARADAT'S INDEX TO SMITH-HAZEL SELECTION INDEX

The equivalence of the two methods can be proved algebraically. Baradat's index can be written as:

$$a'C'V^{-1}p \ldots \text{(AII-1)}.$$

The b values in the classic Smith-Hazel index are:

$$b = V^{-1}Ca \ldots \text{(AII-2)}.$$

Substituting the above expression for the b values in the Smith-Hazel index gives:

$$(V^{-1}Ca)'p \ldots \text{(AII-3)}.$$

The two indexes are equivalent because:

$$(V^{-1}Ca)' = a'C'V^{-1} \ldots \text{(AII-4)}.$$

In Baradat's index the 'b' values are never calculated directly, but $C'V^{-1}p$ is equivalent to the index value of the Smith-Hazel index when the 'a' values in the Smith-hazel index are equal to one.

APPENDIX III.

DERIVATION AND CALCULATION OF BLUP PREDICTIONS

The derivation of BLUP is based on the minimization of the prediction error under the constraint that the estimate is unbiased. The BLUP fromula will be derived for a single trait and then the modifications required for multiple trait selection will be described. The BLUP predictions are based on the linear model:

$$y = X\beta + Za + e \quad \ldots(\text{AIII-1}),$$

where y is a vector of record values which can be measured in an experiment,

X is a design matrix of the fixed effects for each record,

β is a vector of unknown fixed effects,

Z is a design matrix of the random genetic effects for each record,

a is a vector of the unknown genetic effects, and

e is a vector of residual effects.

The derivation and application of BLUP is based on three mathematical procedures:

1. The derivation of basic BLUP formula.
2. The proof of the equivalence of BLUP to selection index when fixed effects are maximum likelihood estimates.
3. The use of mixed model methodology to determine BLUP

solutions.

I. Derivation of basic BLUP formula

The prediction of an unbiased genetic value within a linear combination of fixed effects is:

$$g = \mathbf{k}'\beta + v \quad \ldots\text{(AIII-2)},$$

where g is the value of the genotypes within the $\mathbf{k}'\beta$ fixed effects combination,

k' is a line vector of the linear combination of fixed effects,

β is a column vector of the fixed effects, and

v is the genetic value of the genotype.

The objective is to predict g with a linear combination of y values :

$$\hat{g} = \mathbf{t}'\mathbf{y} \quad \ldots\text{(AIII-3)},$$

where $\hat{g}$ is the prediction of g,

t is a line vector of the coefficients which when applied to the y vector gives $\hat{g}$, and

y is a vector of record scores or subclass means.

The best linear prediction part of the definition requires that the variance of the prediction error be minimized:

$$\mathrm{Var}(t'y - g) = \mathbf{t}'\mathbf{V}\mathbf{t} - 2\mathbf{t}'\mathbf{c} + \mathbf{V} \quad \ldots\text{(AIII-4)},$$

where $\mathbf{V}$ is the variance-covariance matrix of the y values, and c is a vector of the covariance between the genetic value of the genotype within the fixed effects and the y records.

The unbiased prediction part of the definition requires that:

$$t'X = k \quad ...(AIII-5).$$

The minization equation which satisfies the above constraint is:

$$F = (t'Vt - 2t'c + V) + 2\lambda(t'X - k) \quad ...(AIII-6),$$

where λ is a vector of lagrangian multipliers.

Equating the partial derivatives of F with respect to t and λ to zero gives the equations:

$$\begin{bmatrix} V & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} t \\ \lambda \end{bmatrix} = \begin{bmatrix} c \\ k \end{bmatrix} \quad ...(AIII-7).$$

The solutions for t can be calculated by inverting the coefficients matrix. Multiplying the t vector by the records vector y gives an unbiased predictor of g.

II. The equivalence of BLUP to selection indices when fixed effects are maximum likelihood estimates.

The equivalence of BLUP to selection indices has been proved by both Goldberger (1964) and Henderson (1963). Henderson (1984) gave a much simpler proof based on the manipulation of the basic BLUP equations (equation (AIII-7)). The first set of equations are solved for t:

$$t = -V^{-1} + V^{-1}c \quad ...(AIII\text{-}8).$$

Substituting the t value in the second set of equations gives:

$$XV^{-1}X\lambda = -k + XV^{-1}c \quad ...(AIII\text{-}9).$$

Rearranging terms gives the value for λ:

$$\lambda = -(X'V^{-1}X)^{-1}k + (X'V^{-1}X)^{-1}X'V^{-1}c \quad ...(AIII\text{-}10).$$

Substituting the value λ back into the first set of equations gives:

$$t = -V^{-1}X(X'V^{-1}X)^{-1}k - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}c + V^{-1}c \quad ...(AIII\text{-}11).$$

Multiplying throughout by y gives:

$$t'y = k'(X'V^{-1}X)^{-1}X'V^{-1}y + c'V^{-1}(y - X(X'V^{-1}X)^{-1}X'V^{-1}y)$$
$$...(AIII\text{-}12).$$

Substituting the maximum likelihood estimate $\hat{\beta} = (X'V^{-1}X)^{-1}XV^{-1}y$ gives:

$$t'y = k'\hat{\beta} + c'V^{-1}(y - X\hat{\beta}) \quad ...(AIII\text{-}13).$$

Note that the coefficients for BLUP are not the same as the b coefficients in the Smith-Hazel index. The $c'V^{-1}(y - X\hat{\beta})$ is the same as the selection index proposed by Bulmer with the records adjusted with maximum likelihood estimates of the fixed effects. The t values of BLUP are the coefficients which when applied to a vector of y values give an unbiased estimate for a desired linear combination of fixed effects and the selection index of the records adjusted with maximum likelihood estimates.

III. Mixed model methodology to determine BLUP solutions

In the previous section BLUP solutions could be computed after β has been estimated from the formula :

$$\beta = (X'V^{-1}X)^{-}XV^{-1}y \quad \ldots(\text{AIII-14}).$$

V^{-1} is not diagonal and if it is large will require considerable computational effort to invert it. Thus the solution is of limited application for solving practical problems. Mixed model equations of the maximum likelihood estimates of β and u can be calculated by maximizing the joint distribution giving the equations below (Searle 1972):

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ u \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} (\text{AIII-15})$$

where R^{-1} is the inverted matrix of the variance of the error effects vector, and

D^{-1} is the inverted matrix of the variance of the genetic effects vector.

D can be decomposed into:

$$D = A\sigma_a^2 \quad \ldots(\text{AIII-16}),$$

where A is the numerator relationship matrix for the individuals represented in the genetic effects vector, and

σ_a^2 is the additive genetic variance.

Multiplying throughout by σ_e^2 gives:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} + \mathbf{D}^{-1}\sigma_e^2 \end{bmatrix} \begin{bmatrix} \beta \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Z'y} \end{bmatrix} \quad \text{(AIII-17)}$$

Henderson (1976) and Quaas (1976) derived a method of inverting $\mathbf{D}^{-1}$ based on mendelian principles that allowed very large matrices to be inverted. The solutions for β and u can be found by iterative methods which allow equations with up to 100,000 unknowns to be solved (Schaeffer and Kennedy 1985). If the coefficient matrix is inverted then the number of variables which can be solved for will depend on the computer hardware and will generally be limited to under 300.

BLUP multiple trait selection is based upon modification of equation (AIII-15) (Henderson and Quaas 1976). The linear model is expanded to include multiple traits. In the case of selection for three traits, the linear model would be:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & & \\ & \mathbf{X}_2 & \\ & & \mathbf{X}_3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & & \\ & \mathbf{Z}_2 & \\ & & \mathbf{Z}_3 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix} \quad \text{(AIII-18)}$$

where $\mathbf{y}_1\mathbf{y}_2\mathbf{y}_3$ are the record vectors for traits one two and three,

$\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$ are the design matrices of the fixed effects of trait one, two and three for each record,

$\beta_1\beta_2\beta_3$ are the column vectors of the unknown fixed effects for trait one, two and three,

$\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3$ are the design matrices of the unknown random effects of trait one, two and three for each record,

$\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3$ are the column vectors of the unknown genetic effects for traits one, two and three,and

$\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3$ are the column vectors of the residuals for trait one, two and three.

The logic is the same as in the single trait BLUP. The mixed model solutions are more complex. The **X** and **Z** matrices of equation (AIII-15) are replaced by the partitioned matrices in equation (AIII-18). Equation (AIII-15) cannot be simplified because **R** is no longer diagonal. R is a series of diagonal submatrices:

$$\begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \mathbf{R}_{23} \\ \mathbf{R}_{31} & \mathbf{R}_{32} & \mathbf{R}_{33} \end{bmatrix} \quad \text{...(AIII-19)},$$

where $\mathbf{R}_{11} \ldots \mathbf{R}_{33}$ are diagonal submatrices with the diagonal elements equal to variance-covariance between the residual effects for traits one, two and three.

The D matrix becomes:

$$\begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \mathbf{D}_{13} \\ \mathbf{D}_{21} & \mathbf{D}_{22} & \mathbf{D}_{23} \\ \mathbf{D}_{31} & \mathbf{D}_{32} & \mathbf{D}_{33} \end{bmatrix} \quad \text{...(AIII-20)},$$

where $\mathbf{D}_{11} \ldots \mathbf{D}_{33}$ are the product of the relationship matrix **A** and the genetic variance-covariance between traits one, two and three.

Clearly the problem is a more complex computational problem than single trait BLUP. Henderson and Quaas (1976) considered the single animal model, which meant that the **Z** matrix was an identity matrix. Even when the model is restricted to the single animal model there is still a problem with the size of the model and the computational effort required to solve the equations (Quaas and Pollak 1980).

IV. Literature cited

Goldberger, A.S. 1962 Best linear unbiased prediction in the generalized linear regression model. American Statistical Association Journal 57:369-375.

Henderson, C.R. 1963 Selection index and expected genetic advance. Statistical Genetics and Plant Breeding, National Research Council Publication No. 982. National Academy of Science, Washington, D.C. pp. 141-163.

Henderson, C.R. 1976 A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.

Henderson, C.R. 1984 Applications of linear models in animal breeding. University of Guelph. Guelph, Ontario 462 pages.

Henderson, C.R. and R.L. Quaas. 1976 Multiple trait evaluation using relatives records. Journal of Animal Science 43:188-197.

Quaas, R.L. 1976 Computing the diagonal elements and the inverse of a large numerator relationship matrix. Biometrics 32:949-953.

Schaeffer, L.R. and B.W. Kennedy. 1986 Computing strategies for solving mixed model equations. Journal of Dairy Science 69:575-579.

Searle, S.R. 1972 Linear models. John Wiley and Sons, New York. 532 pages.

Quaas, R.L. and E.J. Pollak. 1980 Mixed model methodology for farm and ranch beef cattle testing programs. Journal of Animal Science 51:1277-1287.

APPENDIX IV.

CONSTRUCTION OF BARADAT'S SELECTION INDEX FOR PROVENANCE-PROGENY MODEL

The selection index is a weighted vector of predicted additive genetic values. The predicted additive genetic values are calculated from the formula:

$$\hat{v} = C'V^{-1}pv \ldots (AIV-1).$$

where $\hat{v}$ is a vector of predicted additive genetic value,

- C is a matrix of the covariance between the additive genetic values to be predicted and the phenotypic predictors,
- V^{-1} is a matrix of the variance of the phenotypic predictors, and
- pv is a vector of phenotypic predictors.

To simplify the explanation, the construction of the index will be described for the prediction of the individual's within-provenance additive genetic value. The prediction of the provenance additive genetic value is achieved by modifying the C matrix. This modification will be described in the construction of the C matrix Although we are only concerned with the individual trees additive genetic value it has several predictors. Consider a population of trees grown from seed collected from several mother trees within each of several provenances. The trees developed from the seed of a single tree will all belong to the same half sib family. The trees developed from the seed of each tree in a provenance will all have a common provenance phenotypic value. Thus an individual's phenotypic value can be decomposed into three phenotypic values that can predict its breeding or additive genetic value.

If the additive genetic values of more than one trait are to be predicted then the predictor becomes a linear combination of the phenotypic values of each trait. Assuming that the objective of the selection is to improve a linear combination of three traits the prediction formula expressed in matrix form would be:

$$\left[\hat{v}\right] = \left[C_1 C_2 C_3\right] \begin{bmatrix} V_1^{-1} & & \\ & V_2^{-1} & \\ & & V_3^{-1} \end{bmatrix} \begin{bmatrix} pv_1 \\ pv_2 \\ pv_3 \end{bmatrix} \text{(AIV-2)}$$

where $\hat{v}$ is a 3×1 sub-vector of the individual's predicted additive genetic value for the three traits,

C_1 is a 3×3 sub-matrix of the covariances between the individual's provenance phenotypic values and the individual's additive genetic values for the three traits,

C_2 is a 3×3 sub-matrix of the covariances between the individual's family phenotypic values and the individual's additive genetic values for the three traits,

C_3 is a 3×3 sub-matrix of the covariances between the individual's within family phenotypic values and the individual's additive genetic values for the three traits,

V_1 is a 3×3 sub-matrix of the variance-covariance of the individual's provenance phenotypic values for the three traits,

V_2 is a 3×3 sub-matrix of the variance-covariance of the individual's family phenotypic values for the three traits,

V_3 is a 3×3 sub-matrix of the variance-covariance of the individual's within family phenotypic values for the three traits,

pv_1 is a 3×1 sub-vector of the individual's provenance phenotypic values,

pv_2 is a 3×1 sub-vector of the individual's family phenotypic values, and

pv_3 is a 3×1 sub-vector of the individual's within family phenotypic values.

The phenotypic values are determined from the linear model:

$$Y_{ijk} = \mu + PROV_i + FAM_{j(i)} + e_{k(ij)} \quad \ldots (AIV\text{-}3),$$

where Y_{ijk} is the record value of the k^{th} individual in the j^{th} family belonging to the i^{th} provenance,

μ is the overall mean,

$PROV_i$ is the phenotypic value of the i^{th} provenance,

$FAM_{j(i)}$ is the phenotypic value of the j^{th} family in the i^{th} provenance,

$e_{k(ij)}$ is the within family phenotypic value of the k^{th} individual in the j^{th} family of the i^{th}

provenance.

The three phenotypic predictors and the individual's phenotypic value can be calculated from the basic data:

$$\hat{IND}_{ijk} = Y_{ijk} - \hat{Y}_{...} \quad ...(AIV-4)$$

$$\hat{PROV}_{i} = \hat{Y}_{i..} - \hat{Y}_{...} \quad ...(AIV-5)$$

$$\hat{FAM}_{j(i)} = \hat{Y}_{ij.} - \hat{Y}_{i..} \quad ...(AIV-6)$$

$$\hat{e}_{k(ij)} = Y_{ijk} - \hat{Y}_{ij.} \quad ...(AIV-7)$$

where $\hat{Y}_{...}$ is the average of all the record values;

$\hat{Y}_{i..}$ is the average of all the record values in the i^{th} provenance, and

$\hat{Y}_{ij.}$ is the average of all the record values in the j^{th} family in the i^{th} provenance.

I. Construction of the C matrix

The covariance between the additive genetic values of the individual and the estimated phenotypic predictors is determined by expressing the expectation value of the estimated phenotypic predictors in terms of the estimated phenotypic values of the individuals which contribute to the predictor. If the estimates of the individual phenotypic values are unbias the covariance between the individual phenotypic values which form the phenotypic predictor and the individuals additive genetic values can then be expressed in terms of additive genetic variance. Two important genetic relationships are required for further understanding of the discussion.

1. The expected covariance between an individual's additive genetic value and its estimated phenotypic value expressed as $Y_{ijk} - \hat{Y}_{...}$ is the additive genetic variance (VA).
2. The expected covariance between an individual's additive genetic value and the estimated phenotypic value of the individuals in the same half sib family is a quarter of the additive genetic variance (.25VA).

A. Covariance between the individual's additive genetic value and its provenance phenotypic value

Consider the covariance between the additive genetic value of the ijk^th individual and the estimated phenotypic value of the individual's provenance:

$$\mathrm{Cov}[a_{ijk},(\hat{Y}_{i..} - \hat{Y}_{...})] \quad ...\text{(AIV-8)}.$$

where a_{ijk} is the additive genetic value of the ijkth individual.

There are $n_{i..}$ individual phenotypic values in $(\hat{Y}_{i..} - \hat{Y}_{...})$. One of these phenotypic values is from the ijkth individual. The covariance of this value with a_{ijk} will be equal to VA. There are $n_{ij.} - 1$ individual phenotypic values in $(\hat{Y}_{i..} - \hat{Y}_{...})$ which are half sib relatives of the ijkth individual. Their covariance with a_{ijk} is $(n_{ij.} - 1).25VA$. The expectation of the covariance between the estimated provenance phenotypic value and the additive genetic value of the ijkth individual can be rewritten as:

$$\frac{.25(n_{ij.} - 1) + 1}{n_{i..}} VA \text{ ...(AIV-9).}$$

B. Covariance between the individual's additive genetic value and its family phenotypic value

The estimated family phenotypic value for the j^{th} family nested within the i^{th} provenance must be rewritten so that all deviations are expressed as deviations from the grand mean:

$$\hat{Y}_{ij.} - \hat{Y}_{i..} = (\hat{Y}_{ij.} - \hat{Y}_{...}) - (\hat{Y}_{i..} - \hat{Y}_{...}) \text{ (AIV-10).}$$

The covariance between the individual's additive genetic value a_{ijk} and the family phenotypic predictor is:

$$\mathrm{Cov}[a_{ijk}, (\hat{Y}_{ij.} - \hat{Y}_{...}) - (\hat{Y}_{i..} - \hat{Y}_{...})] \text{ ...(AIV-11)}$$

$$\mathrm{Cov}[a_{ijk}, (\hat{Y}_{ij.} - \hat{Y}_{...})] - \mathrm{Cov}[a_{ijk}, (\hat{Y}_{i..} - \hat{Y}_{...})] \text{ ...(AIV-12).}$$

One of the $n_{ij.}$ inidividuals in the $\hat{Y}_{ij.} - \hat{Y}_{...}$ is $Y_{ijk} - \hat{Y}_{...}$. The expected covariance between the $Y_{ijk} - \hat{Y}_{...}$ and the additive genetic value a_{ijk} is equal to VA. The rest of the individuals in the j^{th} family nested in the i^{th} provenance are half sib relatives if the ijk^{th} individual. Their covariance with a_{ijk} is $(n_{ij.} - 1).25VA$. Therefore the expectation of the covariance term is:

$$\frac{.25(n_{ij.} - 1) + 1}{n_{ij.}} VA \text{ ...(AIV-13).}$$

The expectation of the second covariance term in formula AIV-12 has been derived previously. Substituting the expected values in equation AIV-12 gives the expected covariance between the individual's additive genetic value a_{ijk} and its family phenotypic value:

$$\frac{.25(n_{ij.} - 1) + 1}{n_{ij.}}VA - \frac{.25(n_{ij.} - 1) + 1}{n_{i..}}VA \quad ...(AIV-14).$$

C. Covariance between the individual's additive genetic value and its within-family phenotypic value

The covariance between the estimated phenotypic value $\hat{e}_{ijk}$ and the additive genetic value of the ijkth individual is: $\text{cov}[a_{ijk},((Y_{ijk} - \hat{Y}_{...}) - (\hat{Y}_{ij.} - \hat{Y}_{...}))]$

$$= \text{cov}[a_{ijk},(Y_{ijk} - \hat{Y}_{...})] - \text{cov}[a_{ijk},(\hat{Y}_{ij.} - \hat{Y}_{...})] \quad ...(AIV-15)$$

The first term in the above equation is equal to VA. The expectation of the second term has been determined previously:

$$\frac{VA + ((n_{ij.} - 1).25VA)}{n_{ij.}} \quad ...(AIV-16).$$

Writing the first and second terms together gives:

$$VA - \frac{VA + ((n_{ij.} - 1).25VA)}{n_{ij.}} \quad ...(AIV-17).$$

Rearranging terms gives:

$$\frac{(n_{ij.} - 1)(1 - .25)}{n_{ij.}} VA \quad ...(AIV-18).$$

The C_1, C_2 and C_3 sub-matrices can now be calculated. The model is complicated by spliting the additive variance into two components:

1. Additive variance among provenances.
2. Additive variance within provenances.

VA is calculated from the family variance estimated in the analysis of variance. The model changes only slightly to include the additive variance among provenances. The only difference being that the covariance matrix between the estimated provenance phenotypic values and the additive genetic values of an individual changes. The among provenance additive genetic variance is added to each element of the C_1 matrix. The additive genetic variance among provenances is equal to the provenance variance as measured in the analysis of variance.

II. Construction of the V matrix

The expectation values of the variance of the estimated phenotypic values can be expressed in terms of variance components in the same way as the expected mean squares are expressed in terms of variance components in analysis of variance.

A. Variance of provenance phenotypic value

Consider the variance of the estimate of a provenance phenotypic value. The expectation of the estimate of the i^{th} provenance phenotypic value expressed in terms of the components of the linear model is:

$$\frac{\left[n_{i..}\mathbf{PROV}_i + \sum_{j=1}^{nf} n_{ij.}\mathbf{FAM}_{j(i)}\right]}{n_{i..}} \quad \text{...(AIV-19).}$$

where $n_{i..}$ is the number of individuals in the i^{th} provenance,

$n_{ij.}$ is the number of individuals in the j^{th} family in the i^{th} provenance, and

nf is the number of families in the i^{th} provenance.

The variance of the i^{th} provenance phenotypic value is:

$$\frac{\left[n_{i..}^{2}\sigma_{PROV}^{2} + \sum_{j=1}^{nf} n_{ij.}^{2}\sigma_{FAM(PROV)}^{2}\right]}{n_{i..}^{2}} + \frac{1}{n_{i..}}\sigma_e^{2} \quad \text{...(AIV-20).}$$

The elements of the variance-covariance sub-matrix $\mathbf{V}_1$ can now be determined. The appropriate covariance terms are substituted for variance when determining the non-diagonal elements of the sub-matrix.

B. Variance of family phenotypic value

The elements of the $\mathbf{V}_2$ sub-matrices are derived in a similar manner to formula AIV-20 to give the formula:

$$\frac{\left[n_{i..}^{2} - \sum_{j=1}^{nf} n_{ij.}^{2}\right]}{n_{i..}^{2}}\hat{\sigma}_{FAM(PROV)}^{2} + \left[\frac{1}{n_{ij.}} - \frac{1}{n_{i..}}\right]\sigma_e^{2} \quad \text{...(AIV-21).}$$

The elements of the variance-covariance sub-matrix $\mathbf{V}_2$ can now be determined. The appropriate covariance terms are substituted for variance when determining the non-diagonal elements of the sub-matrix.

C. Variance of the within-family phenotypic value

The elements of the V_3 sub-matrices are derived in a similar manner to formula AIV-20 to give the formula:

$$\frac{n_{ij.} - 1}{n_{ij.}}\sigma_e^2 \quad \ldots\text{(AIV-22)}.$$

The elements of the variance-covariance sub-matrix V_3 can now be determined. The appropriate covariance terms are substituted for variance when determining the non-diagonal elements of the sub-matrix.

The V^{-1} and C matrices change as the family and provenance of the individual for which the additive value is to be predicted changes. The variance components σ_{PROV}^2, σ_{FAM}^2, σ_{WFAM}^2 and VA remain the same but the values $n_{ij.}$, $n_{i..}$ and nf change.

APPENDIX V.

UANOVA PROGRAM

This appendix includes a brief conceptual discription of the UANOVA program and the way it calculates effects, sum of squares and variance components for type I sum of squares. More complex algorithms have been developed for type III sums of squares, however the basic concepts are the same. For more detailed questions about the UANOVA program the reader is refered to Dr. T. Taerum who developed the program at the University of Alberta. Tests using data sets from standard texts show that UANOVA gives the same results as SAS (Taerum 1987).

I. Calculation of effects

The algorithm for calculating the effects is the basic algorithm on which all UANOVA results are based. Basic input values for the algorithm are:

1. Identfication vectors assigning treatment levels to each cell.
2. The mean value of all the records in each cell refered to as the cell mean.
3. The number of records in each cell.

Consider the data set in Table AV-1. There are two treatments provenance and family. The provenance treatment has three levels and the family treatment has two levels. Height is measured on each individual.

Table AV-1. Basic data for example

Individual	Provenance	Family	Height
1	1	1	18
2	1	1	17
3	1	1	16
4	1	2	11
5	1	2	9
6	1	2	7
7	2	1	13
8	2	1	12
9	2	1	11
10	2	2	2
11	2	2	4
12	2	2	6
13	3	1	13
14	3	1	12
15	3	1	14
16	3	2	6
17	3	2	5
18	3	2	4

The input values for the algorithm for the data set are given in Table AV-2.

Table AV-2. Input values for algorithm

Identification vectors			
Provenance	Family	Number of records	Cell mean
1	1	3	17
1	2	3	9
2	1	3	12
2	2	3	4
3	1	3	13
3	2	3	5

The algorithm for finding the effects for each level of each treatment has five basic steps. Steps 1 and 2 initialize the algorithm. Steps 3, 4 and 5 are the basic iteration procedure which continues until the residuals are insignificant. The basic steps are:

1. Calculation of the overall mean.
2. Subtract overall mean from cell means to give residuals.
3. Calculate effects from residuals weighted by the number of individuals in each cell.
4. Subtract effects from residuals to give new residuals.
5. Check residuals against zero value. If significantly different go to step 3 and calculate effec[illegible]or next treatment in the linear model.

When residuals are not significantly different from zero the effects of each level of each treatment are given by summing all the effects estimates for each iteration for that level of treatment. The algorithm will be demonstrated using the data set in Table AV-1.

Step 1:

Calculation of overall mean:

Sum = 18 + 17 + 16 + 11 + 9 + 7 + 13 + 12 + 11 + 2 + 4 + 6 + 13 + 12 + 14 + 6 + 5 + 4 = 180

overall mean = 180 + 18 = 10

Step 2:

Subtract overall mean from cell means:

Cell identified by		Cell mean	Overall mean	Residual
Provenance	Family			
1	1	17	10	7
1	2	9	10	-1
2	1	12	10	2
2	2	4	10	-6
3	1	13	10	3
3	2	5	10	-5

Step 3:

Calculate effects from residuals:

Provenance 1 first estimate

3 × 7 = 21

3 × -1 = -3

18 ÷ 6 = 3

The residual is multipied by the number of records in the cell. In the example above 3 × 7 and 3 × -1. The total is then divided by the number of individuals in the cells which recieved treatment level provenance 1. In this case 18 ÷ 6.

Provenance 2 first estimate

3 × 2 = 6

3 × -6 = -18

-12 ÷ 6 = -2

Provenance 3 first estimate

3 × 3 = 9

3 × -5 = -15

-6 ÷ 6 = -1

Step 4:

Subtract effects from residuals:

Cell identified by Provenance	Family	Old residual	Effect	New residual
1	1	7	3	4
1	2	-1	3	-4
2	1	2	-2	4
2	2	-6	-2	-4
3	1	3	-1	4
3	2	-5	-1	-4

Step 5:

Check residuals against zero value, if all residuals greater then zero go to step 3 to estimate family effects.

Step 3:

Calculates effects from new residuals:

Family 1 first estimate

$3 \times 4 = 12$

$3 \times 4 = 12$

$3 \times 4 = 12$

$36 \div 9 = 4$

Family 2 first estimate

$3 \times -4 = -12$

$3 \times -4 = -12$

$3 \times -4 = -12$

$-36 \div 9 = -4$

Step 4:

Subtract effects from residuals:

Cell identified by Provenance	Family	Old residual	Effect	New residual
1	1	4	4	0
1	2	-4	-4	0
2	1	4	4	0
2	2	-4	-4	0
3	1	4	4	0
3	2	-4	-4	0

Step 5:

Check residuals against zero value. All residuals not significantly different from zero therefore end iterations.

In the above example the residuals converge to zero after one round of iterations, this is because the design is balanced. When designs are unbalanced the residuals will not converge to zero after the first round of iterations. A treatment effect estimate is the sum of all the estimates calculated at each iteration before convergence is reached.

II. Calculating the sum of squares from effects

The sum of squares are calculated from the effects estimates and the number of records in each treatment level. In the example the provenance effects were estimated to be:

Provenance 1 = +3

Provenance 2 = -2

Provenance 3 = -1

There are six records in each provenance. The provenance sum of squares is the sum of the sum of squares for each provenance level. The sum of squares for a given provenance level is:

$(\text{effect})^2 \times$ number of records in treatment level.

In the example the sum of squares for provenance is:

$+3^2 \times 6 = 54$

$-2^2 \times 6 = 24$

$-1^2 \times 6 = 6$

84

The sum of squares for family is:

$+4^2 \times 9 = 144$

$-4^2 \times 9 = 144$

288

The error sum of squares is calculated from the cell means and the record values. The residual or error effect is the difference between the call mean and the record value. The error effects in the example are:

1, 0, -1, 2, 0, -2, 1, 0, -1, -2, 0, 2, 0, -1, 1, 1, 0, -1

The error sum of squares is the sum of the square of these values, that is 24.

This method of calculating sum of squares is very simple. It is based on an accurate estimate of the effects. A summary table of the sum of squares and means squares for the sample data is given in Table AV-3.

Table AV-3. Summary of sum of squares and mean squares for sample data

Treatment	df	Sum of squares	Mean squares
Provenance	2	84	42
Family	3	288	96
Residual	12	24	2

III. Estimating variance components

The variance components are estimated by the synthesis method. The synthesis method uses the design matrix of the linear model. The design matrix of the data in the example is:

μ	P1	P2	P3	F1	F2	F3	F4	F5	F6
1	1	0	0	1	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
1	1	0	0	0	1	0	0	0	0
1	1	0	0	0	1	0	0	0	0
1	1	0	0	0	1	0	0	0	0
1	0	1	0	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0	0
1	0	1	0	0	0	0	1	0	0
1	0	1	0	0	0	0	1	0	0
1	0	1	0	0	0	0	1	0	0
1	0	0	1	0	0	0	0	1	0
1	0	0	1	0	0	0	0	1	0
1	0	0	1	0	0	0	0	1	0
1	0	0	1	0	0	0	0	0	1
1	0	0	1	0	0	0	0	0	1
1	0	0	1	0	0	0	0	0	1

where μ is the mean vector of the design matrix,

P1 is the first provenance effect vector of the design matrix, P2 is the second provenance effect vector of the design matrix,

P3 is the third provenance effect vector of the design matrix,

F1 is the first family effect vector of the design matrix,

F2 is the second family effect vector of the design matrix,

F3 is the third family effect vector of the design matrix,

F4 is the fourth family effect vector of the design matrix,

F5 is the fifth family effect vector of the design matrix, and

F6 is the sixth family effect vector of the design matrix.

Note that each family in each provenance has a separate vector in the design matrix. The synthesis method gives the expectation of the sum of squares in terms of the variance components by passing the design matrix vectors associated with each treatment level through the sum of squares algorithm. Coefficients for variance components in the expectation are calculated by summing the sum of squares algorithm output for each design matrix vector associated with the treatment.

Using the UANOVA sum of squares algorithm there are two stages:

1. Estimation of effects using the vector of the design matrix as input.
2. Calculation of the sum of squares from the effects estimated from the design matrix.

1. Estimation of effects

The procedure will be demonstrated using the sample data. Consider the design matrix vector for provenance 1. Applying the 5 steps of the effects algorithm described earlier gives:

Step 1:

Calculation of the overall mean:

Sum = 1 + 1 + 1 + 1 + 1 + 1 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 = 6

Overall mean = 6 + 18 = .33

Step 2:

Subtract overall mean from cell means:

Cell identified by				
Provenance	Family	Cell mean	Overall mean	Residual
1	1	1	.33	.66
1	2	1	.33	.66
2	1	0	.33	-.33
2	2	0	.33	-.33
3	1	0	.33	-.33
3	2	0	.33	-.33

Step 3:

Calculate effects from residuals:

Provenance 1 first estimate

3 × .66 = 2.0

3 × .66 = 2.0

4.0 ÷ 6 = .66

Provenance 2 first estimate

3 × -.33 = -1.0

3 × -.33 = -1.0

-2.0 ÷ 6 = -.33

Provenance 3 first estimate

3 × -.33 = -1.0

3 × -.33 = -1.0

-2.0 ÷ 6 = -.33

Step 4:

Subtract effects from residuals:

Cell identified by Provenance	Family	Old residual	Effect	New residual
1	1	.66	.66	0
1	2	.66	.66	0
2	1	-.33	-.33	0
2	2	-.33	-.33	0
3	1	-.33	-.33	0
3	2	-.33	-.33	0

Step 5:

Check residuals against zero value. All residuals are not significantly different from zero therefore end iterations.

The same algorithm is applied to the P2 and P3 design matrix vectors. The effects estimated by the algorithm for P1, P2 and P3 vectors are:

Provenance	Family	P1	P2	P3
1	1	.66	-.33	-.33
1	2	.66	-.33	-.33
2	1	-.33	.66	-.33
2	2	-.33	.66	-.33
3	1	-.33	-.33	.66
3	2	-.33	-.33	.66

The residual of each provenance effect vector is zero. When the family design matrix vectors are passed through the algorithm there are provenance and family effects. The provenance effects are used to calculate the coefficient of the family variance component in the provenance sum of squares. The family effects are used to calculate the coefficient of the family variance component in the family sum of squares. The provenance effects estimated by the algorithm for the F1 to F6 vectors are:

Provenance	Family	F1	F2	F3	F4	F5	F6
1	1	.33	.33	-.17	-.17	-.17	-.17
1	2	.33	.33	-.17	-.17	-.17	-.17
2	1	-.17	-.17	.33	.33	-.17	-.17
2	2	-.17	-.17	.33	.33	-.17	-.17
3	1	-.17	-.17	-.17	-.17	.33	.33
3	2	-.17	-.17	-.17	-.17	.33	.33

The Family effects for the design matrix vectors F1 to F6 are:

Provenance	Family	F1	F2	F3	F4	F5	F6
1	1	.5	-.5	0	0	0	0
1	2	-.5	.5	0	0	0	0
2	1	0	0	.5	-.5	0	0
2	2	0	0	-.5	.5	0	0
3	1	0	0	0	0	.5	-.5
3	2	0	0	0	0	-.5	.5

2. Calculation of sum of squares

There are two sums of squares calculations one for the provenance sum of squares and one for the family sum of squares. The coefficients for the provenance variance components for the provenance sum of squares expectation is equal to:

provenance sum of squares P1 + provenance sum of squares P2 + provenance sum of squares P3.

The provenance sum of squares for P1 is:

$+.66^2 \times 6 = 2.66$

$-.33^2 \times 6 = .66$

$-.33^2 \times 6 = .66$

4.00

The provenance sum of squares for P2 is:

$-.33^2 \times 6 = .66$

$+.66^2 \times 6 = 2.66$

$-.33^2 \times 6 = .66$

4.00

The provenance sum of squares for P3 is:

$-.33^2 \times 6 = .66$

$-.33^2 \times 6 = .66$

$+.66^2 \times 6 = 2.66$

4.00

The coefficient for the provenance variance components in the provenance sum of squares is:

$4.0 + 4.0 + 4.0 = 12.0$

The coefficient for the family variance component in the provenance sum of squares is equal to the sum of the sum of squares provenance calculated from the provenance effects estimated from the the vectors F1 to F6. The provenance sum of squares for F1 are:

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

1.00

The provenance sum of squares for F2 are:

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

1.00

The provenance sum of squares for F3 are:

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

1.00

The provenance sum of squares for F4 are:

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

1.00

The provenance sum of squares for F5 are:

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

1.00

The provenance sum of squares for F6 are:

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$-.17^2 \times 3 = .08$

$+.33^2 \times 3 = .33$

$+.33^2 \times 3 = .33$

1.00

The coefficient for the family variance components of the provenance sum of squares is:

$1.0 + 1.0 + 1.0 + 1.0 + 1.0 + 1.0 = 6.0$

The coefficient of the provenance variance component in the expectation value of the family sum of squares is zero because the provenance vectors in the design matirx gave no family effects. The coefficient of the family variance component for the family sum of squares is the sum of the sum of squares of the family effects calculated from the design vectors F1 to F6.

The family sum of squares for F1 are:

$+.5^2 \times 3 = 0.75$

$-.5^2 \times 3 = 0.75$

1.50

The family sum of squares for F2 are:

$-.5^2 \times 3 = 0.75$

$+.5^2 \times 3 = 0.75$

1.50

The family sum of squares for F3 are:

$+.5^2 \times 3 = 0.75$

$-.5^2 \times 3 = 0.75$

1.50

The family sum of squares for F4 are:

$-.5^2 \times 3 = 0.75$

$+.5^2 \times 3 = 0.75$

1.50

The family sum of squares for F5 are:

$+.5^2 \times 3 = 0.75$

$-.5^2 \times 3 = 0.75$

1.50

The family sum of squares for F6 are:

$-.5^2 \times 3 = 0.75$

$+.5^2 \times 3 = 0.75$

1.50

The coefficient for the family variance component in the family sum of squares is:

$1.5 + 1.5 + 1.5 + 1.5 + 1.5 + 1.5 = 9.0$

The coefficients for the expectation value of the mean squares are obtained by dividing the coefficients for the sum of squares by the degrees of freedom which where used to calculate the mean squares. The coefficient for error or residual variance in the mean squares is always one. A summary table of the expected mean squares is given in Table AV-4.

Table AV-4. Summary of expected mean squares calculated by synthesis method

Treatments	Expected mean squares
Provenance	$\sigma_e^2 + 3\sigma_{fam(prov)}^2 + 6\sigma_{prov}^2$
Family(Provenance)	$\sigma_e^2 + 3\sigma_{fam(prov)}^2$
Residual	σ_e^2

The UANOVA program is efficient in terms of computer time. The effects and sum of squares algorithm are quicker than other methods used in standard statistical packages (Taerum 1987). The synthesis method has not been extensively used in computer programs because of the effort required to compute sum of squares in conventional programs. The efficient sum of squares algorithm in UANOVA removes the computational difficulties. Type III or unique sum of squares can be calculated using the same basic principle as described above. Effects are least squares effects and are orthgonalised before sum of squares are calculated.

The UANOVA program was used in this thesis because it had the following advantages:

1. The program calculates effects which are the phenotypic value inputs for selection index calculations.
2. Large data sets can be handled efficiently by the program.
3. The results produced by the program are the same as the results produced by SPSS and SAS (Taerum 1987).
4. Program output includes variance-covariance matrices of dependant variables in the linear model.

IV. Literature cited

Taerum, T. 1987 Efficient algorithms for analysis of variance. Internal Publication, University of Alberta. 27 pages.

APPENDIX VI.

INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER TWO.

Input values and selection index calculations can be represented in a written format. BLUP calculations require the inversion of a 189 × 189 matrix which is impossible to represent in any presentable form. Numerical examples of the BLUP calculation will use a small data set which is not related to the data used in Chapter two but has the same structure.

I. Selection index calculations

The basic input values for the selection index calculations in Chapter two are the phenotypic effects and the variance components. Provenance, family and within family phenotypic effects are given in Tables AVI-1, AVI-2 and AVI-3.

Table AVI-1. Provenance phenotypic effects

Provenance	Height
1	36.615
2	-11.930
3	-21.201
4	29.631
5	-16.732
6	24.527
7	-23.796

Table AVI-2. Family phenotypic effects

Provenance	Family	Height
1	1	-36.506
1	2	161.307
1	3	-24.139
1	4	-26.602
1	5	14.411
1	6	-88.854
2	1	-71.558
2	2	58.730
2	3	-29.248
2	4	-14.582
2	5	16.641
2	6	48.762
3	1	111.958
3	2	-77.860
3	3	-97.426
3	4	187.453
3	5	-39.966
3	6	-85.029
4	1	-10.269
4	2	115.938
4	3	-198.921
4	4	-24.196
4	5	88.574
4	6	29.756
5	1	17.633
5	2	-16.723
5	3	150.850
5	4	-101.294
5	5	-67.628
6	1	-39.827
6	2	157.718
6	3	52.936
6	4	-94.241
6	5	-52.720
6	6	-51.064
7	1	-79.717
7	2	90.690
7	3	88.695
7	4	-119.644
7	5	-102.169
7	6	23.739

7

Table AVI-3. Within-family phenotypic effects

Individual	Provenance	Family	Height
1	1	1	-100.046
2	1	1	99.954
3	1	1	-0.046
4	1	1	-0.046
5	1	2	62.454
6	1	2	-7.546
7	1	2	-7.546
8	1	2	-47.546
9	1	3	-52.546
10	1	3	-62.546
11	1	3	57.454
12	1	3	57.454
13	1	4	139.954
14	1	4	-20.046
15	1	4	-60.046
16	1	4	-60.046
17	1	5	19.954
18	1	5	69.954
19	1	5	-50.046
20	1	5	-40.046
21	1	6	62.454
22	1	6	-147.546
23	1	6	62.454
24	1	6	22.454
25	2	1	69.954
26	2	1	-0.046
27	2	1	-70.046
28	2	2	-0.046
29	2	2	9.954
30	2	2	-10.046
31	2	3	47.454
32	2	3	-182.546
33	2	3	-42.546
34	2	3	177.454
35	2	4	52.454
36	2	4	112.454
37	2	4	-57.546
38	2	4	-107.546
39	2	5	-17.546
40	2	5	112.454
41	2	5	-177.546
42	2	5	82.454
43	2	6	-20.046
44	2	6	9.954
45	2	6	9.954
46	3	1	118.884
47	3	1	28.884
48	3	1	38.884

Table AVI-3. Within-family phenotypic effects (continued)

Individual	Provenance	Family	Height
49	3	1	-1.116
50	3	1	4.752
51	3	1	-45.248
52	3	1	-145.248
53	3	2	-20.461
54	3	2	29.539
55	3	2	169.539
56	3	2	29.539
57	3	2	55.407
58	3	2	-34.593
59	3	2	-64.593
60	3	2	-164.593
61	3	3	-29.952
62	3	3	40.048
63	3	3	30.048
64	3	3	50.048
65	3	3	-164.084
66	3	3	-44.084
67	3	3	-54.084
68	3	3	35.916
69	3	3	135.916
70	3	4	-215.461
71	3	4	-55.461
72	3	4	-75.461
73	3	4	-95.461
74	3	4	0.407
75	3	4	80.407
76	3	4	220.407
77	3	4	140.407
78	3	5	10.312
79	3	5	-29.688
80	3	5	40.312
81	3	5	-69.688
82	3	5	96.180
83	3	5	-203.820
84	3	5	156.180
85	3	6	-0.599
86	3	6	9.401
87	3	6	195.269
88	3	6	-124.731
89	3	6	-214.731
90	3	6	135.269
91	4	1	89.992
92	4	1	-50.008
93	4	1	39.992
94	4	1	-50.008
95	4	1	-30.008
96	4	2	-67.508

Table AVI-3. Within-family phenotypic effects (continued)

Individual	Provenance	Family	Height
97	4	2	82.492
98	4	2	2.492
99	4	2	-17.508
100	4	3	66.659
101	4	3	-23.341
102	4	3	-43.341
103	4	4	-70.008
104	4	4	-60.008
105	4	4	-0.008
106	4	4	129.992
107	4	5	59.992
108	4	5	-60.008
109	4	6	79.992
110	4	6	-80.008
111	5	1	-80.098
112	5	1	19.902
113	5	1	-40.098
114	5	1	19.902
115	5	1	-114.847
116	5	1	55.153
117	5	1	85.153
118	5	1	55.153
119	5	2	-301.794
120	5	2	28.206
121	5	2	63.456
122	5	2	43.456
123	5	2	83.456
124	5	2	83.456
125	5	3	45.736
126	5	3	-74.264
127	5	3	-84.264
128	5	3	10.986
129	5	3	80.986
130	5	3	20.986
131	5	4	21.599
132	5	4	91.599
133	5	4	81.599
134	5	4	51.599
135	5	4	-193.151
136	5	4	-53.151
137	5	5	-223.793
138	5	5	0.956
139	5	5	20.956
140	5	5	130.956
141	5	5	70.956
142	6	1	-117.937
143	6	1	-47.937
144	6	1	32.063

Table AVI-3. Within-family phenotypic effects (continued)

Individual	Provenance	Family	Height
145	6	1	92.063
146	6	1	42.063
147	6	2	-14.937
148	6	2	-4.937
149	6	2	55.063
150	6	2	-34.937
151	6	3	90.063
152	6	3	130.063
153	6	3	-79.937
154	6	3	-139.937
155	6	4	76.729
156	6	4	-53.271
157	6	4	-23.271
158	6	5	25.063
159	6	5	85.063
160	6	5	-34.937
161	6	5	-74.937
162	6	6	93.396
163	6	6	-106.604
164	6	6	13.396
165	7	1	10.063
166	7	1	-9.937
167	7	2	30.063
168	7	2	40.063
169	7	2	-69.937
170	7	3	172.063
171	7	3	32.063
172	7	3	2.063
173	7	3	-7.937
174	7	3	-197.937
175	7	4	-139.937
176	7	4	140.063
177	7	5	12.563
178	7	5	62.563
179	7	5	-67.437
180	7	5	-7.437
181	7	6	-22.437
182	7	6	-92.437
183	7	6	57.563
184	7	6	57.563

The variance components are given in Table AVI-4.

Table AVI-4. Height variance components.

	Provenance	Family	Within-family
Variance	178.419	2153.608	11123.047

In addition calculation requires a knowledge of the number of trees in each family and provenance. These values can be calculated from the number of individuals in each family (Table AVI-5).

Table AVI-5. Number of trees in each family.

Provenance	Family					
	1	2	3	4	5	6
1	4	4	4	4	4	4
2	3	3	4	4	4	3
3	7	8	9	8	7	6
4	5	4	3	4	2	2
5	8	6	6	6	5	
6	5	4	4	3	4	3
7	2	3	5	2	4	4

The variances and phenotypic effects are the input values from which genetic values are predicted.

A. Prediction of genetic values

The formulae for the calculation of the C and V matrices for the three variable case are derived in Appendix IV. The calculation of the elements of the V matrix and the c vector for the single variable case is exactly the same as for the three variable case. The first individual in the second family of the first provenance will be used as an example to demonstrate the calculation method.

i. Calculation of the c vector

The c vector is a 3 by 1 vector. There are two types of additive values:

1. Additive genetic values among provenances
2. Additive genetic values within provenances

The selection index which predicts both additive values adds the provenance additive genetic variance to the first entry in the c vector (Appendix IV). To calculate the elements of the c vector the among provenance additive genetic variance and the within-provenance additve genetic variance must be calculated. The among provenance additive genetic variance is equal to the provenance variance component (Table AVI-4). The within provenance additive variance is equal to four times the family variance that is:

$$4 \times 2153.608 = 8614.432$$

The first element in the c vector is calculated from formula AIV-9. The formula requires the input of the number of trees in the provenance and the number of trees in the families within the provenance for provenance 1 (Table AVI-5). Substituting these values in formula AIV-9 gives:

$$(((0.25 \times (4.0 - 1.0)) + 1.0) \div 24) \times VA$$

Substituting the additive genetic variance components for height gives:

$$(((0.25 \times (4.0 - 1.0)) + 1.0) \div 24) \times 8614.432$$

$$= 628.136$$

To complete the first element of the c vector the provenance additive variance must be added to give:

$$628.136 + 178.419 = 805.56$$

The second element in the c vector is given by substituting the correct values in formula AIV-14. Substituting the number of trees in the family and the number of trees in the provenance the formula becomes:

$$((((0.25 \times (4.0 - 1.0)) + 1.0) \div 4.0) \times VA)$$
$$- ((((0.25 \times (4.0 - 1.0)) + 1.0) \div 24.0) \times VA)$$

Substituting the within-provenance additive genetic variance for height gives:

$$((((0.25 \times (4.0 - 1.0)) + 1.0) \div 4.0) \times 8614.43)$$
$$- ((((0.25 \times (4.0 - 1.0)) + 1.0) \div 24.0) \times 8614.43)$$

$$= 3140.45$$

The third element in the c vector is given by substituting the correct values in formula AIV-18. Substituting the number of trees in the family and the number of trees in the provenance the formula becomes:

$$(((4.0 - 1.0) \times (1.0 - 0.25)) \div 4.0) \times VA)$$

Substituting the within-provenance additive genetic variance for height gives:

$$(((4.0 - 1.0) \times (1.0 - 0.25)) + 4.0) \times 8614.432$$

$$= 4845.618$$

In summary the complete c vector is:

$$\begin{matrix} 806.56 \\ 3140.45 \\ 4845.62 \end{matrix}$$

ii. Calculation of the V matrix

The V matrix is a 3 by 3 diagonal matrix. The first diagonal element is given by substituting the correct values in formula AIV-20. Substituting the number of trees per family and the number of trees in the provenance gives:

$$((576\sigma_{PROV}^{2} + 96\sigma_{FAM(PROV)} \div 576) + ((1 + 24)\sigma_{e}^{2}$$

Substituting the variance components from Table AVI-4 gives:

$$(((576 \times 178.42) + (96 \times 2,153.61)) \div 576) + ((1 + 24) \times 11123.05)$$

$$= 1000.815$$

The second diagonal element is given by substituting the correct values in formula AIV-21. Substituting the number of trees per family and the number of trees in the provenance gives:

$$(((576 - 96) \div 576) \times \sigma_{FAM(PROV)}^{2}) + (((1 \div 4) - (1 \div 24)) \times \sigma_{e}^{2})$$

Substituting the values for the variance components from Table AVI-4 gives:

$$(((576 - 96) \div 576) \times 2153.61) + (((1 \div 4) - (1 \div 24)) \times 11123.05)$$

$$= 4111.977$$

The third diagonal element is given by substituting the correct values in formula AIV-22. Substituting the number of trees per family and the number of trees in the provenance gives:

$$((4 - 1) \div 4) \times \sigma_{e}^{2}$$

Substituting the value for the variance component from Table AVI-4 gives:

$$((4 - 1) \div 4) \times 11123.05$$

$$= 8342.29$$

In summary the complete V matrix is:

$$\begin{array}{rrr} 1000.82 & 0.00 & 0.00 \\ 0.00 & 4111.98 & 0.00 \\ 0.00 & 0.00 & 8342.29 \end{array}$$

iii. Prediction of genetic values from the c vector, V matrix and phenotypic predictors

Formula II-2 is used to predict genetic values. The c vector and V matrix for the first individual in the second family of the first provenance have beeen calculated above. The phenotypic predictors can be read off Tables AVI-1, AVI-2 and AVI-3. The p vector is:

$$\begin{matrix} 36.615 \\ 161.308 \\ 62.453 \end{matrix}$$

The V matrix must be inverted for formula II-2. The inverse of the V matrix calculated above is:

$$\begin{matrix} 0.000999 & 0.000000 & 0.000000 \\ 0.000000 & 0.000243 & 0.000000 \\ 0.000000 & 0.000000 & 0.000120 \end{matrix}$$

Substituting the c vector, V^{-1} matrix and p vector in formula II-2 gives the answer 188.92. The calculation described above was performed with a calculator. The calculations in the main text were performed on a main frame computer. Allowing for rounding errors the predicted value above verifies the value on line 4 of Table II-4 in the main text.

II. BLUP calculations

The data set used to demonstrate the BLUP calculation is given in Table AVI-6.

Table AVI-6. Data set for BLUP calculation example

Block	Provenance	Family	Height
1	1	1	700
1	1	1	750
1	1	2	800
1	1	2	850
1	2	1	900
1	2	1	700
2	2	2	850
2	2	2	750
2	2	2	650
2	3	1	800
2	3	1	850
2	3	2	800
2	3	2	750

The variance components in Table AVI-4. will be used in the demonstration calculation.

The BLUP formula is given in formula II-5. The BLUP calculation can be split into 3 distinct phases:

1. Calculation of the coefficients matrix on the left hand side of the equation.
2. Calculation of the matrix vector products on the right hand side of the equation.
3. Premultiplication of the right side of the equation by the inverse of the coefficients matrix to give the solution to the effects which are to be predicted.

A. Calculation of the coefficients matrix

The basic elements in the coefficients matrix are **X**, Z_1, Z_2, G_1, G_2, s_1 and s_2. The **X** matrix is:

$$\begin{matrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{matrix}$$

Note that the mean is not included this is necessary in order to ensure that the coefficients matrix is of full rank. The consequence of this is that the mean is confounded with each fixed effect. This is normal in BLUP calculations. The Z_1 matrix is:

$$\begin{matrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{matrix}$$

The Z_2 matrix is:

$$\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}$$

The X, Z_1 and Z_2 matrices combine in the coefficients, matrix to produce the matrix:

$$\begin{matrix}
6 & 0 & 4 & 2 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 7 & 0 & 3 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 0 & 4 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 3 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 4 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}$$

To complete the coefficients matrix $G_1^{-1}s_1$ must be added to the $Z_1'Z_1$ part of the coefficients matrix and $G_2^{-1}s_2$ must be added to the $Z_2'Z_2$ part of the matrix. The $Z_1'Z_1$ part of the coefficients matrix is:

$$\begin{matrix}
4 & 0 & 0 \\
0 & 5 & 0 \\
0 & 0 & 4
\end{matrix}$$

The G_1 matrix is:

$$\begin{matrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{matrix}$$

The inverse of the G_2 matrix is the same as the G_2 matrix. The scalar s_1 is given in formulae II-6 and II-7. Substituting the variance components from Table AVI-4 in formula II-7 gives:

$$11123.047 - (3 \times 2153.608)$$

$$= 4663.823$$

The variance of the provenance additive genetic effect is equal to the variance component associated with the provenance effect (Table AVI-4). Substituting the variance components in formula II-6 gives:

$$4663.823 + 178.419$$

$$= 26.1397$$

Multiplying the inverse of the G_1 matrix by this scalar gives:

$$\begin{matrix} 26.1397 & 0.0000 & 0.0000 \\ 0.0000 & 26.1397 & 0.0000 \\ 0.0000 & 0.0000 & 26.1397 \end{matrix}$$

Adding this matrix to the $Z_1'Z_1$ matrix gives:

$$\begin{matrix} 30.1397 & 0.0000 & 0.0000 \\ 0.0000 & 31.1397 & 0.0000 \\ 0.0000 & 0.0000 & 30.1397 \end{matrix}$$

The $Z_2'Z_2$ part of the coefficients matrix is:

$$\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}$$

The G_2 matrix is given on page 203. The G_2^{-1} matrix is given on page 204. The s_2 scalar is given in formula II-8. The σ_{eb}^2 has been calculated previously and the σ_a^2 is four times the family variance. Substituting the variance components in formula II-8 gives:

$$4663.823 \div 8614.432$$

$$= 0.541396$$

Multiplying the G_2^{-1} by this scalar gives the matrix on page 205. Adding this matrix to $Z_2'Z_2$ gives the matrix on page 206. The completed coefficients matrix can be seen on page 207.

1.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.25	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	1.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.25	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	1.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.25	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.25	0.25	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.25	1.00	0.25	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.25	1.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.25	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	1.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.25
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	1.00

G_2

$$\begin{matrix}
1.067 & -0.267 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
-0.267 & 1.067 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 1.067 & -0.267 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & -0.267 & 1.067 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 1.067 & -0.267 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & -0.267 & 1.067 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.111 & -0.222 & -0.222 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.222 & 1.111 & -0.222 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.222 & -0.222 & 1.111 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.067 & -0.267 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.267 & 1.067 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 1.067 & -0.267 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.267 & 1.067
\end{matrix}$$

G_2^{-1} :

```
 0.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
-0.14455   0.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000  -0.14455   0.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000  -0.14455   0.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.60149  -0.12019  -0.12019   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.12019   0.60149  -0.12019   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.12019  -0.12019   0.60149   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.57767  -0.14455   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.14455   0.57767   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.57767  -0.14455
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.14455   0.57767
```

$G_2^{-1} S_2$

```
 1.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
-0.14455   1.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   1.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000  -0.14455   1.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   1.57767  -0.14455   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000  -0.14455   1.57767   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   1.60149  -0.12019  -0.12019   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.12019   1.60149  -0.12019   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.12019  -0.12019   1.60149   0.00000   0.00000   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   1.57767  -0.14455   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.14455   1.57767   0.00000   0.00000
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   1.57767  -0.14455
 0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000   0.00000  -0.14455   1.57767
```

$Z_2'Z_2 + G_2^{-1}S_2$

```
6.00   0.00   4.00   2.00   0.00   1.00   1.00   1.00   1.00   1.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
0.00   7.00   0.00   3.00   4.00   0.00   0.00   0.00   0.00   0.00   0.00   1.00   1.00   1.00   1.00   1.00   1.00   1.00
4.00   0.00  30.14   0.00   0.00   1.00   1.00   1.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
2.00   3.00   0.00  31.14   0.00   0.00   0.00   0.00   0.00   1.00   1.00   1.00   1.00   1.00   0.00   0.00   0.00   0.00
0.00   4.00   0.00   0.00  30.14   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   1.00   1.00   1.00   1.00
1.00   0.00   1.00   0.00   0.00   1.07  -0.27   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
1.00   0.00   1.00   0.00   0.00  -0.27   1.07   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
1.00   0.00   1.00   0.00   0.00   0.00   0.00   1.07  -0.27   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
1.00   0.00   1.00   0.00   0.00   0.00   0.00  -0.27   1.07   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00
1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   1.07  -0.27   0.00   0.00   0.00   0.00   0.00   0.00   0.00
1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00  -0.27   1.07   0.00   0.00   0.00   0.00   0.00   0.00   0.00
0.00   1.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   1.11  -0.22  -0.22   0.00   0.00   0.00   0.00
0.00   1.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00  -0.22   1.11  -0.22   0.00   0.00   0.00   0.00
0.00   1.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00  -0.22  -0.22   1.11   0.00   0.00   0.00   0.00
0.00   1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   1.07  -0.27   0.00   0.00
0.00   1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00  -0.27   1.07   0.00   0.00
0.00   1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   1.07  -0.27
0.00   1.00   0.00   0.00   1.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00   0.00  -0.27   1.07
```

The completed coefficients matrix

B. Calculation of the matrix vector products on the right hand side of the equation

The $\mathbf{X}'y$ vector is the product of the $\mathbf{X}$ matrix and the y vector. The $\mathbf{X}$ matrix has been given in the previous section. The y vector is a vector of record values given in column 4 of Table AVI-6. The $\mathbf{Z}_1'y$ and $\mathbf{Z}_2'y$ are also vector matrix products of the vector y and matrices which have been given previously. The completed vector on the righthand side of formula II-5 is:

$$\begin{array}{r} 4700 \\ 5450 \\ 3100 \\ 3850 \\ 3200 \\ 700 \\ 750 \\ 800 \\ 850 \\ 900 \\ 700 \\ 850 \\ 750 \\ 650 \\ 800 \\ 850 \\ 800 \\ 750 \end{array}$$

C. Premultiplication of the right hand side of the equation by the inverse of the coefficients matrix

Inverting the coefficients matrix and premultiplying the vector on the righthand side of ther equation by the inverse gives a vector:

$$\begin{array}{r} 785.150 \\ 773.100 \\ -1.950 \\ 1.550 \\ -1.200 \end{array}$$

$$\begin{array}{r}
-85.950 \\
-48.450 \\
39.050 \\
76.550 \\
98.350 \\
-51.450 \\
43.700 \\
-31.300 \\
-106.300 \\
49.250 \\
86.750 \\
24.250 \\
-13.250
\end{array}$$

The vector above contains the BLUP estimates for the effect the effect and estimates are given in Table AVI-7.

Table AVI-7. BLUP estimates of effects

Effect	Estimate
Block 1	785.150
Block 2	773.100
Provenance 1	-1.950
Provenance 2	1.550
Provenance 3	-1.200
Tree 1	-85.950
Tree 2	-48.450
Tree 3	39.050
Tree 4	76.550
Tree 5	98.350
Tree 6	-51.450
Tree 7	43.700
Tree 8	-31.300
Tree 9	-106.30
Tree 10	49.250
Tree 11	86.750
Tree 12	24.250
Tree 13	-13.250

III. Conclusion

The reader should be cautioned that the BLUP calculation above is only one example of the many possible BLUP calculations. It is an unusual example as the provenance effects were considered random so that the comparison could be made with selection index scores and rankings. For a comprehensive understanding of BLUP the reader is refered to the literature cited in the Chapters one and two.

APPENDIX VII.

INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER THREE

I. Phenotypic effects

The basic input values for the calculations in Chapter three are the are the phenotypic effects and the variance components. These effects were calculated from the standardized record values. The phenotypic effects for the index based on homogeneous family variance and the index based on heterogeneous family variance are the same. Provenance, family and within family phenotypic effects are given in Tables AVII-1, AVII-2 and AVII-3.

Table AVII-1. Provenance Phenotypic effects

Provenance	Height	Circ.	Lean
1	0.0602	-0.1890	-0.1890
2	0.3074	0.2650	-0.2650
3	-0.0491	0.0784	-0.0784
4	-0.0852	0.0920	0.0920
5	-0.1490	-0.0928	-0.0928
6	-0.1472	-0.2528	-0.2528
7	0.1424	0.1365	0.1365
8	-0.1227	-0.1145	0.1145
9	-0.2106	-0.1749	-0.1749
10	0.1201	0.1309	0.1309
11	0.2030	-0.0862	0.0862
12	-0.1529	0.3659	0.3659

Table AVII-2. Family Phenotypic effects

Provenance	Family	Height	Circ.	Lean
1	1	0.3564	0.1694	-0.0073
1	2	0.9053	0.9630	0.4477
1	3	-0.1623	0.0089	-0.0870
1	4	-0.6079	-0.6762	-0.2034
1	5	0.1416	-0.0016	-0.0294
1	6	-0.4505	-0.1665	-0.0338
2	1	-0.1389	-0.0982	0.1898
2	2	0.3076	0.1887	-0.0258
2	3	0.3790	0.1462	0.1334
2	4	-0.1042	-0.1669	-0.2413
2	5	-0.1457	-0.2208	0.0410
2	6	-0.4261	0.1601	-0.2580
3	1	-0.4401	-0.2659	-0.3120
3	2	0.0252	-0.0120	0.0199
3	3	-0.0332	-0.2174	-0.1008
3	4	-0.2711	-0.1441	0.1644
3	5	0.1465	0.1851	0.0983
3	6	0.5038	0.4524	0.1396
4	1	-0.2278	-0.3887	-0.4424
4	2	0.3326	0.5234	0.1655
4	3	0.2232	0.1934	0.6751
4	4	-0.5961	-0.4686	-0.1914
4	5	-0.2689	-0.4957	-0.2138
4	6	0.3263	0.4610	-0.1725
5	1	-0.0608	0.0222	-0.1039
5	2	0.5321	0.5345	0.4450
5	3	-0.1617	-0.1937	0.1794
5	4	-0.0119	-0.1411	0.2355
5	5	0.0745	0.0056	-0.7095
5	6	-0.3711	-0.2095	-0.0130
6	1	0.0278	-0.1411	-0.6692
6	2	0.6175	0.3553	0.6377
6	3	-0.6011	-0.4338	0.3709
6	4	-0.2904	-0.2296	-0.2310
6	5	0.1402	0.3696	0.2820
6	6	0.2374	0.2180	-0.3380
7	1	-0.0598	0.2857	-0.2398
7	2	-0.2332	-0.2670	-0.1083
7	3	0.0645	-0.1462	0.6443
7	4	0.3869	0.1782	-0.3873
7	5	0.0077	-0.0929	-0.1434
7	6	-0.2443	0.1349	0.2289
8	1	0.0430	-0.1223	-0.5601
8	2	-0.7249	-0.4845	-0.0829
8	3	-0.1391	-0.4091	0.2649
8	4	0.7202	0.6450	-0.0934
8	5	-0.2219	-0.0489	0.2374
8	6	0.1568	0.1252	0.1710

Table AVII-2. Family Phenotypic effects (continued)

Provenance	Family	Height	Circ.	Lean
9	1	0.4029	0.3131	0.0613
9	2	-0.3289	-0.5137	-0.1957
9	3	-0.3043	-0.1361	0.0133
9	4	0.4700	0.5371	0.5488
9	5	-0.1987	-0.0650	0.0776
9	6	-0.0286	-0.1488	-0.5915
10	1	0.1193	0.0754	0.1154
10	2	-0.2454	-0.1652	-0.0666
10	3	0.6230	0.2954	0.0457
10	4	-0.3348	-0.3814	-0.0184
10	5	-0.0903	0.1965	-0.1036
11	1	-0.9115	-0.7635	-0.1666
11	2	0.7837	0.5402	-0.1489
11	3	0.4595	0.4498	0.0733
11	4	-0.1662	0.0310	-0.1016
11	5	-0.0331	-0.1429	0.1462
11	6	0.0435	-0.0551	0.2996
12	1	0.3835	0.6506	-0.2289
12	2	0.5574	0.3734	0.2815
12	3	-0.6644	-0.6870	-0.0779

Table AVII-3. Within-family Phenotypic effects

Individual	Provenance	Family	Height	Circ.	Lean
1	1	1	0.8789	0.9063	0.6188
2	1	1	0.7989	-0.1337	-0.4212
3	1	1	0.1189	1.0063	0.9688
4	1	1	-0.7911	-1.6537	-1.4612
5	1	1	-1.1997	-0.8217	0.4901
6	1	1	0.3103	0.5983	1.0101
7	1	1	-0.4397	0.0383	-0.8999
8	1	1	-0.4397	0.0383	-0.0299
9	1	1	0.8110	0.5738	-0.5552
10	1	1	0.2010	-0.4662	0.6648
11	1	1	-0.2490	-0.0862	-0.3852
12	1	2	-0.2819	-0.4584	-1.1769
13	1	2	-1.0319	-1.2084	0.2031
14	1	2	-0.1319	-0.0784	0.5531
15	1	2	-0.1319	-0.1684	0.0331
16	1	2	-0.4319	-0.6384	0.3831
17	1	2	0.1802	1.0091	0.5891
18	1	2	0.9696	0.4737	-0.0956
19	1	2	0.4396	-0.0063	1.2944
20	1	2	0.4396	1.4237	-0.0956
21	1	2	0.1396	0.2837	-2.8756
22	1	2	-0.7198	-0.6009	0.2491
23	1	2	0.5602	-0.0309	0.9391
24	1	3	0.6470	0.4154	0.1912
25	1	3	1.0270	1.0754	0.0112
26	1	3	0.3470	0.6054	0.7112
27	1	3	0.6590	0.3629	0.0473
28	1	3	-0.4010	0.2729	-1.5127
29	1	3	-0.8510	-1.0571	0.0473
30	1	3	-1.1510	-0.9571	0.9173
31	1	3	-1.0010	-1.2471	0.7473
32	1	3	-0.2216	0.0174	-0.1175
33	1	3	-0.2916	-0.4626	0.2325
34	1	3	0.6184	0.4874	-0.1175
35	1	3	0.6184	0.4874	-1.1575
36	1	4	0.2590	-0.3246	0.5983
37	1	4	-0.9510	-0.9846	0.4183
38	1	4	-0.3410	-0.7046	-0.9717
39	1	4	1.0190	0.7154	-0.2717
40	1	4	-0.0410	0.6254	-0.6217
41	1	4	-1.1710	-0.6146	0.2483
42	1	4	1.2390	1.1854	0.2483
43	1	4	0.2590	0.2454	-0.2717
44	1	4	-0.4890	-0.6571	0.6343
45	1	4	-0.7890	-0.6571	0.6343
46	1	4	-1.1690	-0.7571	0.8043
47	1	4	-1.0090	-0.6571	0.6343
48	1	4	0.6510	0.4729	-1.6257

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
49	1	4	0.1210	0.5729	0.4643
50	1	4	1.6604	1.8274	-1.4404
51	1	4	0.4504	0.5974	-0.5804
52	1	4	0.1504	-0.1626	0.2896
53	1	4	0.1504	-0.7226	0.8096
54	1	5	-0.2595	0.5102	0.4226
55	1	5	0.4905	0.8002	0.4226
56	1	5	0.3405	0.3302	0.5926
57	1	5	0.1905	-0.9098	-0.9674
58	1	5	0.5025	0.5576	0.9786
59	1	5	-0.1775	-0.2924	-0.9314
60	1	5	-0.6275	-0.2924	0.1086
61	1	5	-1.0875	-1.1424	0.8086
62	1	5	0.3119	0.9622	-2.3161
63	1	5	0.6819	0.7722	0.6439
64	1	5	-0.2181	-1.0278	0.8139
65	1	5	-0.1481	-0.2678	-0.5761
66	1	6	0.4615	0.1043	0.7786
67	1	6	0.5315	0.1943	0.0786
68	1	6	0.0815	0.1943	-0.4414
69	1	6	-0.8864	-1.0882	0.8146
70	1	6	0.4230	0.6463	-0.5701
71	1	6	-1.1570	-1.1537	-1.2601
72	1	6	0.4230	0.1763	0.2999
73	1	6	0.1230	0.9263	0.2999
74	2	1	1.1734	0.9489	-0.1712
75	2	1	0.5634	0.7689	0.3488
76	2	1	0.4934	0.9489	0.6988
77	2	1	-0.0297	0.1623	-0.0743
78	2	1	-0.4797	-0.4077	-0.2443
79	2	1	-0.2597	-0.2177	0.4557
80	2	1	1.0303	0.9223	-1.1143
81	2	1	-0.6634	-0.5805	-0.9801
82	2	1	-0.2134	-0.7705	-0.1101
83	2	1	-0.5934	-0.3905	0.9299
84	2	1	0.3166	0.2795	0.7599
85	2	1	-1.1576	-1.1827	1.3222
86	2	1	-0.6276	-0.1427	-0.0678
87	2	1	-0.0276	0.1373	-2.1478
88	2	1	0.4224	0.4273	-0.7578
89	2	1	0.0524	-0.9027	1.1522
90	2	2	0.3548	0.3884	-0.2078
91	2	2	-0.0252	0.0084	-0.5478
92	2	2	-0.4752	-0.5616	0.6622
93	2	2	0.0548	0.1084	0.3222
94	2	2	-1.8283	-1.6281	0.2490
95	2	2	-1.6320	-1.2410	0.7233
96	2	2	1.0080	0.9390	1.2433

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
97	2	2	-0.5720	-0.1910	-1.0167
98	2	2	0.6638	0.1468	-0.4444
99	2	2	0.7438	1.0868	0.9456
100	2	2	1.1938	-0.1432	-0.4444
101	2	2	0.5138	1.0868	-1.4844
102	2	3	-0.4800	-0.3317	-1.4366
103	2	3	-0.4800	-0.6117	0.6434
104	2	3	-0.2500	-0.0417	0.4734
105	2	3	0.6600	0.8083	-0.7466
106	2	3	0.5890	0.3766	0.7468
107	2	3	0.8990	1.8866	-1.8532
108	2	3	-0.6910	-0.3834	0.9168
109	2	3	-1.1410	-0.7634	0.7468
110	2	3	-0.2431	0.0217	-1.6897
111	2	3	0.8169	1.0617	1.0903
112	2	3	-0.0931	-0.8383	0.5703
113	2	3	-0.7267	-1.3912	-0.3455
114	2	3	0.2533	0.5088	0.5245
115	2	3	0.5533	0.4188	0.3545
116	2	3	0.3333	-0.7212	0.0045
117	2	4	-0.8964	-0.3985	-0.2497
118	2	4	-0.8964	-1.0585	-0.7697
119	2	4	-0.3664	-0.7785	0.7903
120	2	4	0.5405	-0.1451	0.5471
121	2	4	1.0368	2.0521	0.8514
122	2	4	-0.1274	-0.4502	-0.8463
123	2	4	-1.1074	-1.0201	0.8937
124	2	4	-0.8874	-0.9301	0.5437
125	2	4	1.0005	1.4649	-1.5429
126	2	4	1.2668	-0.0279	0.6714
127	2	4	0.4368	1.2921	-0.8886
128	2	5	0.8781	0.4176	0.9013
129	2	5	0.1981	-0.3424	0.7213
130	2	5	0.7281	0.3276	0.0313
131	2	5	0.8081	0.5176	-0.3187
132	2	5	0.5850	0.5810	0.1282
133	2	5	-1.1786	0.0282	-2.1676
134	2	5	-0.2686	0.2182	1.4724
135	2	5	-0.3486	-0.6418	-2.1676
136	2	5	-0.1629	0.1760	1.0048
137	2	5	0.2871	0.4560	-1.0852
138	2	5	-0.6129	-0.3940	0.3048
139	2	5	-0.9129	-1.3440	1.1748
140	2	6	1.3073	1.4547	-1.1386
141	2	6	0.5573	-0.6253	0.4214
142	2	6	0.5573	0.6047	0.4214
143	2	6	-1.3958	-1.5119	0.6982
144	2	6	-0.1158	1.2381	-0.8618

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
145	2	6	-1.4995	-1.8747	0.1325
146	2	6	0.7605	1.5353	-0.3875
147	2	6	0.6463	0.9231	-0.5152
148	2	6	-0.8637	-2.0069	0.3548
149	2	6	0.0463	0.2631	0.8748
150	3	1	0.2054	0.1502	0.4222
151	3	1	0.8854	0.8102	-0.2678
152	3	1	0.8730	0.8365	-0.8139
153	3	1	-0.5670	-0.3035	1.0961
154	3	1	-0.7170	-1.0635	1.0961
155	3	1	-0.5670	-0.5835	-0.1139
156	3	1	0.2812	0.7807	-0.3261
157	3	1	-0.2488	-0.0693	-0.8461
158	3	1	-0.7788	-0.3593	0.1939
159	3	1	1.0084	0.2853	-0.2001
160	3	1	-0.6516	-0.8547	-0.3701
161	3	1	-0.0516	-0.0047	-0.8901
162	3	1	0.3284	0.3753	1.0199
163	3	2	0.6160	1.1583	0.3112
164	3	2	-0.7440	-0.5417	0.3112
165	3	2	-0.8940	-1.4917	1.3512
166	3	2	0.1829	-0.2969	0.0636
167	3	2	-0.4171	-0.5769	0.0636
168	3	2	0.8629	1.2231	-0.2864
169	3	2	0.4929	-0.0169	0.0636
170	3	2	0.1829	0.0831	0.2336
171	3	2	-0.1295	0.6695	-1.6925
172	3	2	-0.4295	-0.5605	0.7375
173	3	2	-0.5095	-0.0805	0.3875
174	3	2	0.2587	0.0537	-0.1647
175	3	2	0.3387	0.4237	-0.3447
176	3	2	0.1887	-0.0463	-1.0347
177	3	3	-0.2995	-0.1428	-0.1604
178	3	3	0.3005	0.3272	-1.5504
179	3	3	-0.8295	-0.9928	-0.3304
180	3	3	0.2205	-0.5228	0.1896
181	3	3	0.9275	0.9521	-1.9680
182	3	3	0.7075	1.2421	-0.0580
183	3	3	1.0075	0.6721	-1.7980
184	3	3	0.3975	1.0521	0.2820
185	3	3	-1.4249	-1.2016	0.7858
186	3	3	0.2351	0.2184	-0.4242
187	3	3	0.1551	0.3184	1.1358
188	3	3	-0.1449	-0.4416	-0.2542
189	3	3	0.0851	0.3184	0.9658
190	3	3	0.0232	-0.3074	-0.1164
191	3	3	-1.7068	0.2626	0.7536
192	3	3	-0.6568	-1.6374	1.2736

Table. AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
193	3	3	1.0032	-0.1174	1.2736
194	3	4	-1.0530	-1.6419	1.5023
195	3	4	-0.2930	-0.3219	0.9823
196	3	4	0.2370	-0.2219	-0.5777
197	3	4	-0.7430	-0.0319	0.6323
198	3	4	-0.4960	-0.0771	0.7347
199	3	4	-0.1160	0.0229	-0.1353
200	3	4	-0.3584	-1.1907	-0.5014
201	3	4	1.0716	1.1793	-1.5414
202	3	4	1.6816	2.0393	-1.7214
203	3	4	0.4098	1.0335	-1.7536
204	3	4	0.8698	0.9335	0.6764
205	3	4	-0.4202	-1.1465	0.5064
206	3	4	-0.7902	-0.5765	1.1964
207	3	5	-0.4537	0.2637	-1.2414
208	3	5	0.0663	0.3537	-0.2014
209	3	5	-0.3961	0.1001	-2.8276
210	3	5	-0.0961	-0.1899	1.5124
211	3	5	0.6539	0.6601	0.3024
212	3	5	-0.3079	0.5142	0.2602
213	3	5	0.6721	1.1742	-0.2598
214	3	5	-1.5179	-2.1358	0.9602
215	3	5	0.4521	-0.8158	1.1302
216	3	5	-0.1807	-0.3611	0.0462
217	3	5	-0.0307	0.0189	1.0862
218	3	5	1.2493	0.5889	-1.1638
219	3	5	-0.1107	-0.1711	0.3962
220	3	6	-0.5060	1.2614	-1.4720
221	3	6	-0.7740	-0.6286	-0.7820
222	3	6	0.3560	0.7914	-0.2620
223	3	6	0.5060	0.7914	0.0880
224	3	6	-0.1470	-0.9538	0.0104
225	3	6	0.0130	-0.3838	-1.1996
226	3	6	-0.1470	0.1862	-0.5096
227	3	6	-0.4470	-1.6138	0.5304
228	3	6	0.2206	0.8726	1.3842
229	3	6	0.9706	0.3026	0.3442
230	3	6	-0.0794	0.0226	-0.0058
231	3	6	-0.4413	-0.6032	0.4820
232	3	6	-0.2213	0.3368	0.1320
233	3	6	-0.2213	-1.7432	0.1320
234	3	6	-0.0940	1.3614	1.1280
235	4	1	1.0456	1.4194	0.0249
236	4	1	0.1356	-0.2806	0.1949
237	4	1	0.8956	1.3294	-1.1951
238	4	1	-0.4644	-0.1906	-0.6751
239	4	1	-1.0644	-0.6606	0.7149
240	4	1	1.3681	1.2437	0.2419

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
241	4	1	-0.5219	0.1037	0.0719
242	4	1	0.2381	-0.5563	-0.4481
243	4	1	0.2381	0.8637	0.0719
244	4	1	-1.2726	-1.1175	0.6333
245	4	1	-0.2239	-1.1723	-0.3376
246	4	1	-0.3739	-0.9823	0.7024
247	4	2	-0.2635	0.5051	-0.1191
248	4	2	0.1065	-0.0649	-1.3391
249	4	2	1.0865	1.5451	1.0909
250	4	2	-0.7453	-1.0380	1.0125
251	4	2	-0.7453	-1.1280	-0.2075
252	4	2	-1.9553	-1.9780	-0.0375
253	4	2	-0.6428	-0.1637	-0.6805
254	4	2	0.7872	0.3063	0.1895
255	4	2	1.1665	1.2551	-1.1591
256	4	2	0.6252	0.3503	1.6900
257	4	2	0.7052	0.6303	0.4700
258	4	2	-0.1248	-0.2197	-0.9100
259	4	3	0.5139	1.2040	-0.6456
260	4	3	-0.4661	-0.4060	-0.1256
261	4	3	-1.0879	-1.1791	0.8260
262	4	3	0.4221	0.0509	0.4760
263	4	3	1.3221	0.9109	-1.4340
264	4	3	0.4221	-0.5191	-0.9040
265	4	3	-0.7654	-0.7848	0.1830
266	4	3	-1.7454	-1.6448	2.6130
267	4	3	0.4439	-0.4060	1.2644
268	4	3	-0.1561	0.1640	-0.6456
269	4	3	-1.3661	-0.7860	-0.4756
270	4	3	1.7926	2.1992	-0.9165
271	4	3	0.7326	0.1092	0.8135
272	4	3	0.5026	0.4892	0.6435
273	4	3	0.4326	1.4392	-0.5765
274	4	3	-0.9974	-0.8408	-1.0965
275	4	4	-0.6904	-0.9684	1.2470
276	4	4	1.0396	1.4016	-2.5730
277	4	4	-0.0904	-0.1184	1.0770
278	4	4	0.4879	0.5285	-0.0514
279	4	4	0.1179	0.0585	-0.0514
280	4	4	-0.0321	-0.1315	1.1686
281	4	4	0.3604	-0.2171	0.8656
282	4	4	-0.6896	-0.6871	-2.0844
283	4	4	-1.3116	-0.4531	0.6361
284	4	4	0.8084	0.5869	-0.2339
285	4	5	-0.7194	-0.4611	1.2295
286	4	5	0.3406	-0.2811	-0.5005
287	4	5	0.4106	0.6689	0.8895
288	4	5	0.7906	0.6689	-1.0205

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
289	4	5	0.5388	1.5058	-1.7989
290	4	5	0.8388	-0.1042	0.4511
291	4	5	0.5388	0.5558	-0.9389
292	4	5	-0.8687	-1.2299	1.1981
293	4	5	-0.8687	-0.7599	-0.5419
294	4	5	0.7913	0.5701	-0.3719
295	4	5	-0.3507	-0.3359	-0.0814
296	4	5	0.0193	-0.0459	0.9586
297	4	5	-0.9607	-0.5159	-0.4314
298	4	5	-0.5007	-0.2359	0.9586
299	4	6	1.0966	1.9932	-2.5519
300	4	6	0.0466	-0.2868	0.2281
301	4	6	-1.0134	-1.3268	-0.2919
302	4	6	-0.1251	0.2701	0.4897
303	4	6	-0.4351	-0.6799	0.6697
304	4	6	0.0949	0.6501	-0.5503
305	4	6	-0.7351	-0.1099	0.8397
306	4	6	0.6474	0.5644	0.3667
307	4	6	0.1174	-1.0456	0.5367
308	4	6	1.0274	1.5144	-2.0633
309	4	6	-0.3334	-0.7568	0.9281
310	4	6	-0.2646	-0.4316	1.6972
311	4	6	-0.7946	-0.9116	0.8272
312	4	6	0.3354	0.3284	-1.4328
313	4	6	0.3354	0.2284	0.3072
314	5	1	0.6278	0.7297	0.1818
315	5	1	0.0178	0.3497	-0.5182
316	5	1	0.1778	-0.1303	0.3518
317	5	1	-0.5822	-0.7903	0.1818
318	5	1	1.0942	1.2028	-0.0816
319	5	1	1.4742	1.9628	0.6084
320	5	1	1.3242	0.6428	-0.6016
321	5	1	-0.7860	-1.0868	-0.5506
322	5	1	-0.9360	-0.5168	0.1394
323	5	1	-0.5660	-0.4268	0.3194
324	5	1	-1.8460	-1.9368	-0.0306
325	5	2	-0.4231	-0.2692	0.9030
326	5	2	0.4069	0.7708	0.2130
327	5	2	-0.7331	-0.5492	0.9030
328	5	2	0.0433	0.2140	1.3397
329	5	2	-0.5667	0.5940	0.1197
330	5	2	0.7933	0.8740	-1.9603
331	5	2	0.6433	-0.4460	-0.4003
332	5	2	0.6433	0.9740	-2.1403
333	5	2	-0.3369	-0.3756	0.3406
334	5	2	-0.4069	-1.4156	1.5606
335	5	2	0.0431	0.1944	-0.1794
336	5	2	-0.1069	-0.5656	-0.6994

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
337	5	3	-0.4008	-0.9536	0.9871
338	5	3	0.3492	0.9464	-3.5229
339	5	3	0.8792	0.2764	0.8171
340	5	3	-0.7808	-1.0436	1.1571
341	5	3	0.7292	1.0364	-0.0529
342	5	3	0.5156	-0.0004	0.8938
343	5	3	-1.5944	-1.4204	0.5538
344	5	3	-0.2444	-0.1904	-0.6662
345	5	3	-1.2244	-1.2304	0.5538
346	5	3	0.5954	1.8800	-1.4852
347	5	3	-0.3146	-0.0200	-1.4852
348	5	3	0.5954	-0.6800	1.1248
349	5	3	0.8954	1.4000	1.1248
350	5	4	1.0258	1.2711	0.6927
351	5	4	0.7958	1.3611	0.1727
352	5	4	-0.4042	-0.6289	-0.6973
353	5	4	0.0458	0.0311	-0.1773
354	5	4	-0.9178	-1.0957	0.2594
355	5	4	-0.3178	-0.7157	0.0794
356	5	4	0.8122	0.3243	0.9494
357	5	4	-0.8478	-0.6257	0.2594
358	5	4	0.2120	0.3047	-0.2097
359	5	4	-0.1580	-0.8353	-1.0797
360	5	4	-0.0880	-0.2653	-0.7297
361	5	4	-0.1580	0.8747	0.4803
362	5	5	0.1095	-0.2070	0.0958
363	5	5	-0.3405	-0.0170	0.0958
364	5	5	-0.3405	-0.3070	0.0958
365	5	5	-0.4905	-0.3070	0.0958
366	5	5	0.2759	-0.7738	-0.6875
367	5	5	1.1059	1.0262	0.1825
368	5	5	0.7259	0.2762	0.3525
369	5	5	0.2759	1.0262	0.1825
370	5	5	-2.5143	-1.9234	-0.1166
371	5	5	0.0557	-0.9834	0.4034
372	5	5	0.2757	0.2566	0.0534
373	5	5	0.2757	0.4466	0.4034
374	5	5	0.5857	1.4866	-1.1566
375	5	6	0.1041	0.0094	0.2795
376	5	6	0.1041	-0.0905	-1.1105
377	5	6	-0.4259	-0.6506	-0.0705
378	5	6	0.2541	0.3895	0.7995
379	5	6	0.8803	1.3230	0.9371
380	5	6	1.3303	1.8930	-1.3229
381	5	6	0.2003	-0.1970	0.0671
382	5	6	-1.3195	-1.4074	0.1862
383	5	6	-0.1895	0.2026	0.0162
384	5	6	-0.6395	-1.0274	-0.3338

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
385	5	6	-0.0395	-0.4574	0.0162
386	5	6	-0.2595	0.0126	0.5362
387	6	1	0.6292	0.7609	0.0231
388	6	1	0.0992	-0.3791	0.2031
389	6	1	-0.4308	-0.4691	0.5431
390	6	1	0.2492	0.7609	-2.2269
391	6	1	-0.6306	-0.1960	0.2274
392	6	1	-1.3106	-0.8560	0.0574
393	6	1	-0.0306	0.0940	0.2274
394	6	1	0.9650	1.2309	0.2231
395	6	1	-0.0950	-0.1891	0.5731
396	6	1	0.5850	-0.5691	0.2231
397	6	1	-0.0950	-0.1891	-0.2969
398	6	1	0.0650	0.0009	0.2231
399	6	2	0.1056	-0.1200	-2.3083
400	6	2	0.1456	0.8300	-1.0983
401	6	2	-1.3244	-1.0700	0.2917
402	6	2	-0.1642	-0.6968	-0.5440
403	6	2	-0.3942	-0.5968	0.6660
404	6	2	-0.6942	-1.1668	1.1860
405	6	2	0.0715	-0.4099	1.3617
406	6	2	1.2015	1.9601	1.1917
407	6	2	0.6015	0.7301	-0.0283
408	6	2	0.4515	0.5401	-0.7183
409	6	3	-0.9289	-0.9354	0.0307
410	6	3	0.8111	0.3846	-1.6993
411	6	3	0.5011	0.2946	-0.1393
412	6	3	-0.0189	0.1046	0.7307
413	6	3	-0.3989	0.0046	1.4207
414	6	3	0.7513	1.9077	-0.4651
415	6	3	1.0613	0.2877	0.4049
416	6	3	-0.0787	-0.0823	0.9249
417	6	3	-0.0631	-0.4654	-0.9794
418	6	3	-0.7431	-0.6554	0.9306
419	6	3	-0.8931	-0.8454	-1.1594
420	6	4	-0.1904	-0.3903	-0.7443
421	6	4	-0.4204	-0.0203	-0.0543
422	6	4	-1.0204	-0.8703	0.2957
423	6	4	-0.1702	-0.5871	-0.0200
424	6	4	0.6598	0.5529	-1.2400
425	6	4	0.2798	0.0729	-0.3700
426	6	4	-0.9202	-0.6771	0.5000
427	6	4	-0.0845	0.0798	1.0157
428	6	4	-0.0045	0.2698	-0.0243
429	6	4	0.4455	-0.0202	0.4957
430	6	4	1.4255	1.5898	0.1457
431	6	5	-0.6183	-0.8093	0.2973
432	6	5	-0.6983	-0.4293	-0.0527

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
433	6	5	-0.7683	-0.5193	0.8173
434	6	5	-0.6983	-0.8093	0.9873
435	6	5	-0.8980	-1.0861	0.8416
436	6	5	0.7620	0.4239	-1.4084
437	6	5	0.6120	0.3339	-0.3684
438	6	5	0.6120	-0.2361	1.0216
439	6	5	1.2976	2.7008	-2.6328
440	6	5	0.3976	0.4308	0.4972
441	6	6	-0.8753	-1.0378	-0.3234
442	6	6	-0.2053	-0.3778	-0.1534
443	6	6	0.4249	0.0953	0.5709
444	6	6	0.0549	-0.3847	0.4009
445	6	6	0.0549	0.6653	-0.2991
446	6	6	-0.0251	-0.2847	0.0509
447	6	6	0.8906	1.0422	0.9166
448	6	6	-0.3194	0.2822	-1.1634
449	7	1	0.1486	1.2687	0.7470
450	7	1	0.2286	-0.2513	-0.4730
451	7	1	-0.4514	0.4187	-1.1630
452	7	1	-0.0722	-0.5244	0.4144
453	7	1	-0.2222	-0.8044	0.9344
454	7	1	0.0171	0.1657	0.2751
455	7	1	-0.1329	-0.2143	-0.7649
456	7	1	-0.5129	-1.3543	0.1051
457	7	1	0.9971	1.2957	-0.0749
458	7	2	0.1708	0.4921	-0.5115
459	7	2	0.4008	-0.3579	0.5285
460	7	2	-0.3592	0.2121	-0.8615
461	7	2	-1.1799	-1.2910	0.5459
462	7	2	-0.7999	-1.2910	0.7259
463	7	2	0.8601	0.6990	0.5459
464	7	2	0.1893	0.2392	0.7567
465	7	2	0.4193	0.1492	-0.2833
466	7	2	0.3393	0.8092	-0.4633
467	7	2	-0.0407	0.3392	-0.9833
468	7	3	0.1053	0.3710	-1.1477
469	7	3	0.5553	0.0910	-0.1077
470	7	3	-0.5047	-0.6690	0.0723
471	7	3	-0.2047	-0.7590	0.5923
472	7	3	-1.4662	-0.6419	-0.0495
473	7	3	-0.8662	-0.3519	1.1605
474	7	3	0.0438	0.0281	-1.9595
475	7	3	-0.1062	-0.2619	0.4705
476	7	3	0.3345	0.2879	0.9597
477	7	3	1.0945	1.0479	-0.9503
478	7	3	1.0145	0.8579	0.9597
479	7	4	-1.9992	-2.1329	0.2495
480	7	4	0.4208	0.5271	-1.1405

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
481	7	4	-0.2192	0.8071	0.2495
482	7	4	-0.8899	-0.7960	0.0969
483	7	4	-1.0307	-1.0558	0.4777
484	7	4	-0.1307	-0.4858	-0.0423
485	7	4	0.8493	0.5542	0.8277
486	7	4	0.4793	-0.2958	0.6477
487	7	4	1.0701	0.6340	-1.2931
488	7	4	1.4501	2.2440	-0.0731
489	7	5	-0.7438	-0.7214	0.4934
490	7	5	-0.5138	0.1286	0.8434
491	7	5	-0.5938	0.4186	-0.8966
492	7	5	0.9955	0.9955	-3.8292
493	7	5	-0.4345	-0.3345	0.5108
494	7	5	-0.2045	0.2355	-0.1792
495	7	5	-0.1345	0.1455	0.6908
496	7	5	0.1848	-0.1244	0.0216
497	7	5	0.0248	-1.1644	0.2016
498	7	5	0.3348	-0.2144	1.0716
499	7	5	1.0848	0.6356	1.0716
500	7	6	-0.2636	-0.8511	-0.5231
501	7	6	-0.4236	-1.3312	1.2169
502	7	6	0.3364	-0.0012	1.0469
503	7	6	-1.2443	-0.9343	-1.3657
504	7	6	0.8850	1.5459	-0.6450
505	7	6	0.2750	0.7859	0.3950
506	7	6	0.4350	0.7859	-0.1250
507	8	1	0.0737	-0.2827	-0.0790
508	8	1	0.0037	0.0073	0.0910
509	8	1	-0.1463	-0.5627	-1.1190
510	8	1	-0.5263	0.0973	-0.5990
511	8	1	-0.6664	-0.0701	0.4930
512	8	1	1.2759	-0.1146	0.6066
513	8	1	-0.0141	0.9254	0.6066
514	8	2	-0.8904	-0.2884	0.5723
515	8	2	0.0096	-0.0084	0.7523
516	8	2	-0.2904	-0.0084	0.9223
517	8	2	-0.5104	-0.5784	0.0523
518	8	2	-0.3604	-0.0984	0.7523
519	8	2	0.6295	0.1042	0.4543
520	8	2	-0.1205	-0.5658	-0.0657
521	8	2	0.2495	0.3842	-1.9757
522	8	2	0.6018	0.0597	-1.8621
523	8	2	0.6818	0.9997	0.3979
524	8	3	-1.2577	-0.8449	0.4182
525	8	3	-0.4078	-0.9223	-0.0498
526	8	3	0.7922	0.4977	1.3402
527	8	3	0.3945	0.8332	0.0638
528	8	3	0.6945	0.3632	-3.0562

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
529	8	3	-0.2155	0.0732	1.2838
530	8	4	1.4338	0.7609	0.2301
531	8	4	0.4538	0.4709	-0.6399
532	8	4	-0.2262	-0.1891	-0.4699
533	8	4	-0.0762	-0.1891	-0.2999
534	8	4	0.2437	0.0135	0.2722
535	8	4	1.2937	1.3435	1.1422
536	8	4	0.2437	-0.3665	0.6222
537	8	4	-0.9140	-1.6410	0.3957
538	8	4	-1.1440	-0.1310	-2.7343
539	8	4	-0.7640	-0.1310	0.9157
540	8	4	-0.5440	0.0590	0.5657
541	8	5	1.3949	0.7845	-0.8345
542	8	5	0.1849	-0.0655	0.0255
543	8	5	0.1149	0.2145	-0.1445
544	8	5	-1.1751	-1.2055	0.3755
545	8	5	0.5470	0.5626	-1.0190
546	8	5	0.0970	0.0026	-0.4990
547	8	5	0.5470	0.3726	0.8910
548	8	5	-1.9152	-1.7529	1.4675
549	8	5	0.2048	1.0871	-0.2625
550	8	6	0.1885	-0.4286	0.1449
551	8	6	-0.4915	-0.5286	0.3149
552	8	6	0.4185	0.3314	0.6649
553	8	6	-0.2615	-0.3386	0.6649
554	8	6	0.0207	-0.0804	1.5305
555	8	6	-0.8793	-0.9305	0.8305
556	8	6	-0.3493	-0.1705	-0.2095
557	8	6	0.1707	1.3396	-0.8995
558	8	6	-0.6316	-0.4160	0.7170
559	8	6	0.7284	1.1940	-1.0230
560	8	6	0.4284	-0.0360	-2.5830
561	8	6	0.6584	0.0640	-0.1530
562	9	1	0.0815	0.5897	0.1566
563	9	1	-0.3785	-0.3603	0.5066
564	9	1	-0.3785	0.2997	1.1966
565	9	1	-0.3700	-0.7469	-0.0325
566	9	1	-0.5200	-0.5669	-0.2125
567	9	1	-1.4300	-1.3169	1.0075
568	9	1	1.2522	-0.0218	-0.0086
569	9	1	0.5722	0.6382	0.1614
570	9	1	0.6422	0.0682	-0.5286
571	9	1	0.3422	1.0182	-0.5286
572	9	1	0.5689	0.8897	-0.2793
573	9	1	0.1889	-0.3403	-2.3693
574	9	1	-0.5711	-0.1503	0.9307
575	9	2	1.2642	-0.0016	-0.2826
576	9	2	1.4142	1.5084	0.2474

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
577	9	2	0.6542	-0.0016	0.0674
578	9	2	0.6542	0.4684	0.4174
579	9	2	-0.1674	-0.0182	0.5784
580	9	2	-0.9174	-0.6782	0.5784
581	9	2	-0.7674	-0.6782	0.7484
582	9	2	-0.5052	-0.1431	0.2523
583	9	2	-0.1352	0.2369	-0.6177
584	9	2	0.9248	0.8069	-0.4377
585	9	2	-0.1352	-0.5231	1.1223
586	9	2	0.2416	1.0584	-1.5785
587	9	2	-0.4384	-0.5516	0.3315
588	9	2	-0.6684	-0.3616	-2.0985
589	9	2	-1.4184	-1.1216	0.6715
590	9	3	1.1597	1.3216	-0.6045
591	9	3	-0.2703	-0.0084	0.2655
592	9	3	0.2597	0.4716	-0.6045
593	9	3	-0.3503	-0.5684	0.2655
594	9	3	-0.0419	-0.1150	0.5964
595	9	3	-0.7597	-0.8099	1.1503
596	9	3	-0.2297	-0.2399	-0.7597
597	9	3	-0.3097	-0.3299	0.8003
598	9	3	-0.1597	-0.4299	0.9703
599	9	3	-1.5929	-0.5584	0.6995
600	9	3	-0.6929	-0.3684	0.5295
601	9	3	-0.7629	-0.9284	0.1795
602	9	3	-0.0829	-0.4584	0.1795
603	9	3	0.6671	0.4916	-2.4305
604	9	3	1.6981	0.8350	-0.9636
605	9	3	1.4681	1.6950	-0.2736
606	9	4	-0.8960	-0.4930	1.1625
607	9	4	-0.2160	0.2670	0.9925
608	9	4	-1.4260	-1.5330	1.1625
609	9	4	0.0840	0.5570	-0.9175
610	9	4	-1.6476	-1.3496	-0.9365
611	9	4	-1.1176	-0.6896	-0.7665
612	9	4	-1.4976	-1.7296	-0.0665
613	9	4	-0.7376	-0.3096	1.4935
614	9	4	-0.7854	0.5055	-0.9126
615	9	4	0.4246	0.4155	0.8274
616	9	4	0.2746	-0.0545	0.3074
617	9	4	0.1246	0.9855	-0.9126
618	9	4	1.0214	0.0070	0.0266
619	9	4	1.6314	1.0470	-0.8334
620	9	4	2.6814	1.0470	0.2066
621	9	4	2.0814	1.3270	-0.8334
622	9	5	0.6853	0.5892	-0.0110
623	9	5	0.0753	0.1092	-1.0510
624	9	5	0.4553	0.2992	-0.8710

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
625	9	5	-0.1447	-0.8308	0.1690
626	9	5	-1.0463	-1.1275	0.8499
627	9	5	0.7637	0.7625	1.1899
628	9	5	-0.1141	0.2577	-1.3862
629	9	5	-0.4141	-0.7823	0.0038
630	9	5	0.1159	0.0677	0.1738
631	9	5	-0.7141	-1.0723	0.8738
632	9	5	0.7127	1.8392	-0.0970
633	9	5	-1.5473	-1.1908	0.6030
634	9	5	1.1727	1.0792	-0.4470
635	9	6	-0.3966	0.1964	-0.6830
636	9	6	-0.8466	-0.4636	0.5370
637	9	6	-0.3166	-0.2736	-0.1530
638	9	6	0.0534	-0.1736	0.5370
639	9	6	1.5718	1.0397	-0.6920
640	9	6	-0.7360	-1.1751	0.5518
641	9	6	-0.6560	-0.5051	0.7218
642	9	6	0.9208	0.8864	-0.9389
643	9	6	-1.4892	-1.1036	-0.4189
644	9	6	-2.1692	-1.4836	0.2711
645	9	6	0.4708	-0.0636	0.2711
646	9	6	2.1718	2.2697	-0.5220
647	9	6	1.4218	0.8497	0.5180
648	10	1	-0.3730	-0.4790	-0.4967
649	10	1	0.3770	-0.1990	1.2433
650	10	1	-0.3030	-0.0090	0.2033
651	10	1	0.4570	0.9410	0.5533
652	10	1	1.7034	1.5370	-1.0703
653	10	1	0.0434	0.0270	0.1397
654	10	1	0.0434	1.2570	0.3197
655	10	1	-1.0866	-0.7330	-1.5903
656	10	1	-0.9105	-0.8369	0.3942
657	10	1	-0.1505	-0.5469	-0.8158
658	10	1	-0.6105	-0.7369	0.2242
659	10	1	-0.1505	-0.2669	-0.1258
660	10	1	-0.7774	-0.6086	0.8177
661	10	1	0.5026	0.2514	1.1677
662	10	1	0.7326	0.2514	0.8177
663	10	1	0.5026	0.1514	-1.7823
664	10	2	-0.2374	-0.1493	0.0247
665	10	2	-0.2374	0.0407	1.2347
666	10	2	0.5126	0.9907	-1.0153
667	10	2	-0.6174	-0.5293	0.0247
668	10	2	0.0290	-0.0233	0.4811
669	10	2	1.0190	-0.6833	-1.7689
670	10	2	0.1890	0.4567	0.8311
671	10	2	-1.3210	-1.4433	-0.8989
672	10	2	-2.4248	-1.7272	0.7356

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
673	10	2	0.0652	-0.0272	-0.4744
674	10	2	0.7183	0.3911	1.1590
675	10	2	0.5683	0.4911	0.1190
676	10	2	0.8683	1.2511	-1.0910
677	10	2	0.8683	0.9611	0.6390
678	10	3	-0.2274	0.2323	-0.7296
679	10	3	-0.5374	-0.6177	1.3504
680	10	3	-0.8374	-0.3377	-1.0796
681	10	3	-0.6074	-0.4277	0.6604
682	10	3	0.5151	0.1743	0.8513
683	10	3	-0.3849	-0.5857	0.5013
684	10	3	-0.4649	-0.4957	1.5513
685	10	3	0.6482	0.7826	-1.6752
686	10	3	1.1782	0.4026	0.0648
687	10	3	0.7182	0.8726	-1.4952
688	10	4	0.9872	0.0703	-0.8001
689	10	4	-0.0628	0.2603	1.1099
690	10	4	0.4336	0.1063	-0.8537
691	10	4	1.4836	1.7163	0.1863
692	10	4	0.9536	0.6763	-1.7237
693	10	4	-0.6002	-0.9377	0.0908
694	10	4	-0.0702	-0.2777	0.6108
695	10	4	-0.1502	-0.6577	1.6508
696	10	4	-0.3702	0.1923	0.0908
697	10	4	-1.8271	-0.9994	-0.0058
698	10	4	-0.7771	-0.1494	-0.3558
699	10	5	-0.0645	0.0713	1.1828
700	10	5	0.5119	0.3873	0.2592
701	10	5	0.8919	0.3873	0.6092
702	10	5	1.1119	0.9573	0.7792
703	10	5	-0.2481	-0.0827	-0.9508
704	10	5	0.4319	0.7673	1.1292
705	10	5	0.7319	0.8573	-0.9508
706	10	5	-0.7420	-1.1367	0.1637
707	10	5	-0.5920	-0.2767	0.3337
708	10	5	0.2380	0.4833	-0.3563
709	10	5	-0.2220	-0.8467	1.2037
710	10	5	-2.0489	-1.5684	-3.4028
711	11	1	0.6567	0.5878	0.3680
712	11	1	0.1267	0.4878	0.5480
713	11	1	-0.4033	-0.2722	0.0180
714	11	1	1.0988	1.0346	-1.6827
715	11	1	0.2688	-0.2954	-0.8127
716	11	1	1.5488	1.8846	-0.8127
717	11	1	-0.9941	-0.8093	0.6360
718	11	1	-0.3941	-0.5293	-0.4040
719	11	1	-0.2341	-0.4293	0.9860
720	11	1	-1.6741	-1.6593	1.1560

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
721	11	2	-0.1343	-0.0558	-0.3129
722	11	2	-0.3543	0.1342	-0.4929
723	11	2	-0.7343	0.1342	-0.4929
724	11	2	0.7678	0.7710	-1.1537
725	11	2	0.6878	-0.1790	0.2363
726	11	2	0.1578	-0.2690	-0.1137
727	11	2	-0.1952	-0.6929	0.9950
728	11	2	-0.1952	0.1571	1.3350
729	11	3	-2.6960	-1.4117	-0.0458
730	11	3	-1.7160	-1.1317	0.8242
731	11	3	0.8811	0.7244	-0.5071
732	11	3	0.9511	0.9144	0.0129
733	11	3	-0.0289	-0.7956	-1.0271
734	11	3	0.2011	-0.4156	0.7129
735	11	3	-0.5681	-0.8186	-0.5950
736	11	3	1.0919	1.4514	1.1350
737	11	3	1.0919	0.8814	-0.7750
738	11	3	0.7919	0.6014	0.2650
739	11	4	-0.6168	-0.0154	0.9474
740	11	4	0.0632	0.7446	-2.1726
741	11	4	-0.1047	-0.6085	-0.2334
742	11	4	0.8053	0.4315	0.6366
743	11	4	-0.9347	-0.7985	-0.0634
744	11	4	-0.1047	-0.3285	0.6366
745	11	4	0.3023	0.6676	1.2154
746	11	4	-0.5277	-0.6524	0.8654
747	11	4	0.0023	-0.5624	0.6954
748	11	4	-0.5277	-0.3724	-0.1746
749	11	4	1.6432	1.4946	-2.3526
750	11	5	0.8116	0.2376	-0.1680
751	11	5	0.1316	-0.5224	0.1820
752	11	5	-0.5384	-0.9024	0.8820
753	11	5	-1.2184	-0.9024	0.1820
754	11	5	-0.9984	-0.9924	0.1820
755	11	5	1.2807	1.6906	-3.3700
756	11	5	1.3607	1.8806	1.4900
757	11	5	-0.8293	-0.4894	0.6200
758	11	6	-0.3828	-0.9851	1.3667
759	11	6	-0.6028	-0.0351	-0.5433
760	11	6	-0.0028	-0.2251	0.3267
761	11	6	-0.4707	-0.3482	1.0559
762	11	6	0.5364	0.2778	-1.4853
763	11	6	0.7664	1.2278	-0.9653
764	11	6	0.1564	0.0878	0.2447
765	12	1	-0.2117	-0.7836	0.2331
766	12	1	-0.3617	-0.3036	0.9331
767	12	1	0.2383	-0.5936	1.1031
768	12	1	0.4109	0.6830	-1.3820

Table AVII-3. Within-family Phenotypic effects (continued)

Individual	Provenance	Family	Height	Circ.	Lean
769	12	1	0.4109	-0.2670	0.8780
770	12	1	-0.1091	1.6330	-1.3820
771	12	1	-0.1138	-0.5140	0.8483
772	12	1	-0.2638	0.1460	-1.2317
773	12	2	-0.2452	0.5445	0.0580
774	12	2	-0.0152	-0.1255	-1.3320
775	12	2	0.6074	0.4911	-0.1772
776	12	2	-1.2826	-1.8789	1.3928
777	12	2	0.7574	0.2011	0.6928
778	12	2	0.4574	0.8611	0.1728
779	12	2	0.3827	-0.4259	-0.5468
780	12	2	0.3127	0.9941	0.6632
781	12	2	0.0027	-0.0459	0.8432
782	12	2	-0.9773	-0.6159	-1.7668
783	12	3	0.0803	0.1952	-0.3024
784	12	3	0.5403	1.3252	0.7376
785	12	3	0.7603	1.0452	-1.5224
786	12	3	-0.5974	-0.9013	-0.9118
787	12	3	1.4826	2.2187	1.3482
788	12	3	-1.1772	-1.0981	-1.2276
789	12	3	-0.1972	-0.0581	0.6824
790	12	3	-0.3472	-0.6181	1.2024
791	12	3	-0.1572	0.0419	-0.3576
792	12	3	0.4881	-0.3952	-0.2172
793	12	3	0.7181	0.5548	-1.2572
794	12	3	-1.3219	-1.6252	1.5228
795	12	3	-0.2719	-0.6852	0.3028

II. Variance components

The provenance variance components for the homogeneous and heterogeneous family variance model are the same (Table AVII-4).

Table AVII-4. Provenance variance-covariance components

	Height	Circumference	Lean
Height	0.010	0.003	0.003
Circumference	0.003	0.019	-0.006
Lean	0.003	-0.006	0.011

The family variance-covariance components for the homogeneous family variance model are given in Table AVII-5.

Table AVII-5. Family variance-covariance components for homogeneous family variance model

	Height	Circumference	Lean
Height	0.123	0.100	-0.030
Circumference	0.100	0.090	-0.023
Lean	-0.030	-0.023	0.023

The within-family variance components for the homogeneous family variance model are given in Table AVII-6.

Table AVII-6. Within-family variance-covariance components for homogeneous family variance model

	Height	Circumference	Lean
Height	0.633	0.550	-0.162
Circumference	0.550	0.742	-0.237
Lean	-0.162	-0.237	0.938

The family variance components are different for each provenance in the heterogeneous family variance model (Table AVII-7). The within-family variance components are different for each provenance in the heterogeneous family variance model (Table AVII-8).

Table AVII-7. Family variance-covariance components for heterogeneous family variance model

	Height	Circ.	Lean
Provenance 1			
Height	0.296	0.281	-0.113
Circumference	0.281	0.276	-0.116
Lean	-0.113	-0.116	-0.011
Provenance 2			
Height	0.041	-0.022	-0.020
Circumference	-0.022	-0.037	0.019
Lean	-0.020	0.019	-0.041
Provenance 3			
Height	0.073	0.054	-0.017
Circumference	0.054	0.028	-0.012
Lean	-0.017	-0.012	-0.030
Provenance 4			
Height	0.083	0.112	-0.060
Circumference	0.112	0.150	-0.059
Lean	-0.060	-0.059	0.094
Provenance 5			
Height	0.042	0.030	-0.021
Circumference	0.030	0.005	-0.015
Lean	-0.021	-0.015	0.103
Provenance 6			
Height	0.132	0.089	-0.030
Circumference	0.089	0.064	-0.038
Lean	-0.030	-0.038	0.170
Provenance 7			
Height	-0.005	-0.035	0.030
Circumference	-0.035	-0.033	0.056
Lean	0.030	0.056	0.054
Provenance 8			
Height	0.200	0.160	0.041
Circumference	0.160	0.127	0.030
Lean	0.041	0.030	-0.055
Provenance 9			
Height	0.060	0.078	-0.062
Circumference	0.078	0.097	-0.092
Lean	-0.062	-0.092	0.099
Provenance 10			
Height	0.130	0.069	-0.054
Circumference	0.069	0.044	-0.026
Lean	-0.054	-0.026	-0.068
Provenance 11			
Height	0.263	0.203	0.022
Circumference	0.203	0.147	0.039
Lean	0.022	0.039	-0.093
Provenance 12			
Height	0.512	0.520	-0.068
Circumference	0.520	0.530	0.002
Lean	-0.068	0.002	-0.069

Table AVII-8. Within-family variance-covariance components for heterogeneous family variance model

	Height	Circ.	Lean
Provenance 1			
Height	0.445	0.396	-0.046
Circumference	0.396	0.560	-0.054
Lean	-0.046	-0.054	0.725
Provenance 2			
Height	0.631	0.590	-0.091
Circumference	0.590	0.864	-0.229
Lean	-0.091	-0.229	0.928
Provenance 3			
Height	0.441	0.400	-0.227
Circumference	0.400	0.719	-0.311
Lean	-0.227	-0.311	0.896
Provenance 4			
Height	0.684	0.651	-0.301
Circumference	0.651	0.863	-0.443
Lean	-0.301	-0.443	1.098
Provenance 5			
Height	0.623	0.621	-0.090
Circumference	0.621	0.903	-0.191
Lean	-0.090	-0.191	0.801
Provenance 6			
Height	0.417	0.410	-0.178
Circumference	0.410	0.617	-0.271
Lean	-0.178	-0.271	0.873
Provenance 7			
Height	0.533	0.500	-0.106
Circumference	0.500	0.774	-0.232
Lean	-0.106	-0.232	0.938
Provenance 8			
Height	0.541	0.372	-0.150
Circumference	0.372	0.477	-0.143
Lean	-0.150	-0.143	1.166
Provenance 9			
Height	1.028	0.748	-0.270
Circumference	0.748	0.761	-0.269
Lean	-0.270	-0.269	0.694
Provenance 10			
Height	0.623	0.442	0.111
Circumference	0.442	0.485	0.102
Lean	0.111	0.102	1.036
Provenance 11			
Height	0.867	0.725	-0.373
Circumference	0.725	0.813	-0.406
Lean	-0.373	-0.406	1.141
Provenance 12			
Height	0.465	0.495	-0.007
Circumference	0.495	0.915	-0.223
Lean	-0.007	-0.223	1.232

III. Number of trees in each family

The calculations also require a knowledge of the number of trees in each family and the number of trees in each provenance. These values can be calculated from the number of trees in each family (Table AVII-9).

Table AVII-9. Number of trees in each family

Provenance	Family					
	1	2	3	4	5	6
1	11	12	12	18	12	8
2	16	12	15	11	12	10
3	13	14	17	13	13	15
4	12	12	16	10	14	15
5	11	12	13	12	13	12
6	12	10	11	11	10	8
7	9	10	11	10	11	7
8	7	10	6	11	9	12
9	13	15	16	16	13	13
10	16	14	10	11	12	
11	10	8	10	11	8	7
12	8	10	13			

IV. Calculation of C and V matrices

The C and V matrices are the basic matrices from which the genetic worth is predicted and the prediction parameters such as prediction variance are calculated. The C and V matrices change as the family changes. When family variance components are assumed to be homogeneous the magnitude of the changes depends on the number of trees in a family and the number of trees in a provenance. For the range of the number of trees within a family in the data set there was no significant difference between the C and V matrices among families. In the heterogeneous family variance model there are considerable differences for the C and V matrices for each provenance because of changes in variance components (Tables AVII-7 and AVII-8).

A. Calculation of C and V matrices for homogeneous family variance model

All the values for prediction variance and prediction error variance for the homogeneous family variance model in Chapter three are based on estimates calculated from the first family in provenance 6.

i. Calculation of the elements of the C matrix

The C matrix is a 9 by 3 matrix made up of three submatrices as explained in Appendix AIV. There are two types of additive genetic variance:

1. Additive genetic variance among provenances
2. Additive genetic variance within-provenances

The selection index which predicts both additive values adds the provenance variance to the C_1 matrix (Appendix IV).

To calculate the elements of the C matrix the among provenance and the within-provenance additive variance must be calculated. The among provenance additive variance is equal to the provenance variance components (Table AVII-4). The within-provenance additive variance is equal to four times the family variance (Table AVII-10).

Table AVII-10. Within-provenance variance-covariance components for homogeneous family variance model

	Height	Circumference	Lean
Height	0.492	0.400	-0.120
Circumference	0.400	0.360	-0.092
Lean	-0.120	-0.092	0.092

The elements of the C_1 matrix are calculated from formula AIV-9. Substituting the number of trees in a family and the number of trees in a provenance gives:

$$(((0.25 \times (12.0 - 1.0)) + 1.0) \div 62.0) \times VA$$

Substituting the additive genetic variance components for height from Table AVII-10 gives:

$$(((0.25 \times (12.0 - 1.0)) + 1.0) \div 62.0) \times 0.492$$

$$= 0.030$$

The calculation of the other elements of the C_1 matrix is the same except the appropriate within-provenace additive genetic variance or covariance replaces the height within-provenance additive genetic variance. The within-provenance elements of the C_1 matrix are:

	Height	Circumference	Lean
Height	0.030	0.024	-0.007
Circumference	0.024	0.022	-0.006
Lean	-0.007	-0.006	0.006

To obtain the final C_1 matrix the among provenance additive genetic variance must be addded (Table AVII-4). The final C_1 matrix is:

	Height	Circumference	Lean
Height	0.040	0.027	-0.004
Circumference	0.027	0.041	-0.012
Lean	-0.004	-0.012	0.017

The elements of the C_2 matrix are given by substituting the correct values in formula AIV-14. Substituting the number of trees in the family and the number of trees in the provenance in the formula gives:

$$((((0.25(12.0 - 1.0)) + 1.0) + 12.0) \times VA)$$
$$- (((0.25 \times (12.0 - 1.0)) + 1.0) + 62) \times VA)$$

Substituting the within-provenance additive genetic variance for height from Table AVII-10 gives:

$$((((0.25(12.0 - 1.0)) + 1.0) + 12.0) \times 0.492)$$
$$- ((((0.25 \times (12.0 - 1.0)) + 1.0) + 62) \times 0.492)$$

$$= 0.124$$

The calculation of the other elements of the C_2 matrix are the same except the appropriate within-provenance additive variance or covariance replaces the height within-provenance additive variance. The C_2 matrix is:

	Height	Circumference	Lean
Height	0.124	0.101	-0.030
Circumference	0.101	0.091	-0.023
Lean	-0.030	-0.023	0.023

The elements of the C_3 by substituting the correct values in formula AIV-18. Substituting the values for the number of trees in the family and the number of trees in the provenance gives:

$$(((12.0 - 1.0) \times (1.0 - 0.25)) \div 12.0) \times VA$$

$$= 0.338$$

The calculation of the other elements of the C_3 is the same except the appropriate within-provenance additive variance or covariance repalces the height within-provenance additive variance. The C_3 matrix is:

	Height	Circumference	Lean
Height	0.338	0.275	-0.083
Circumference	0.275	0.248	-0.063
Lean	-0.083	-0.063	0.063

In summary the complete **C** matrix is:

0.040	0.027	-0.004
0.027	0.041	-0.012
-0.004	-0.012	0.017
0.124	0.101	-0.030
0.101	0.091	-0.023
-0.030	-0.023	0.023
0.338	0.275	-0.083
0.275	0.248	-0.063
-0.083	-0.063	0.063

ii. Calculation of elements of V matrix

The formulae for calculating the elements of the **V** matrix are given in Appendix IV (Formulae AIV-20, AIV-21 and AIV-22). There are three block diagonal matrices.

The elements of the first block matrix are given by substituting the correct values in formula AIV-20. Substituting the correct values for the number of trees in the provenance and the number of trees in the family in the formula gives:

$$((3844\sigma_{PROV}^2 + 650\sigma_{FAM(PROV)}^2) \div 3844) + ((1 \div 62) \times \sigma_e^2)$$

Substituting the height variance components gives:

$$(((3844 \times 0.010) + (650 \times 0.123)) \div 3844) + ((1 \div 62) \times 0.633)$$

$$= 0.041$$

The calculation of the other elements of the first block diagonal matrix is the same but the appropriate variance or covariance components replace the height variance component. The first block diagonal matrix of the V matrix is:

	Height	Circumference	Lean
Height	0.041	0.029	-0.005
Circumference	0.029	0.046	-0.014
Lean	-0.005	-0.014	0.030

The elements of the second block matrix are given by substituting the correct number of trees in the family and in the provenance into formula AIV-21. The formula becomes:

$$(((3844 - 650) \div 3844) \times \sigma_{FAM(PROV)}^2) + (((1 \div 12) - (1 \div 62)) \times \sigma_e^2)$$

Substituting the values for height from Tables AVII-5 and AVII-6 gives:

$$(((3844 - 650) \div 3844) \times 0.123) + (((1 \div 12) - (1 \div 62)) \times 0.633)$$

$$= 0.145$$

The calculation of the other elements of the second block diagonal matrix is the same except the appropriate variance or covariance components replace the height variance components. The second block diagonal matrix is:

	Height	Circumference	Lean
Height	0.145	0.120	-0.036
Circumference	0.120	0.125	-0.035
Lean	-0.036	-0.035	0.082

The elements of the third block diagonal matrix are given by substituting the correct values for the number of trees in the family in formula AIV-22. The formula becomes:

$$(11 \div 12) \times \sigma_e^2$$

Substituting the value for height from Table AVII-6 gives:

$$(11 \div 12) \times 0.633$$

$$= 0.580$$

The calculation of the other elements of the third block matrix is the same except the appropriate variance or covariance component replaces the height variance component. The third block diagonal matrix of the V matrix is:

	Height	Circumference	Lean
Height	0.580	0.504	-0.149
Circumference	0.504	0.680	-0.217
Lean	-0.149	-0.217	0.860

In summary the completed V matrix is:

$$\begin{matrix}
0.041 & 0.029 & -0.005 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.029 & 0.046 & -0.014 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
-0.005 & -0.014 & 0.030 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.144 & 0.120 & -0.036 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.120 & 0.125 & -0.035 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & -0.036 & -0.035 & 0.082 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.580 & 0.504 & -0.149 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.504 & 0.680 & -0.217 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.149 & -0.217 & 0.860
\end{matrix}$$

B. Calculation of C and V matrices for heterogeneous family variance model

The C and V matrices for the heterogeneous family variance model are calculated by substituting the correct values in the same formulae used in the homogeneous family variance model. The C and V matrices different for each for each provenance. The family and within-family variance components change for each provenance. Therefore new families were choosen for the calculation of the C and V matrices used in the calculation of the prediction variance, genetic worth variance and efficiency (Table AVII-11).

Table AVII-11. Families from which C and V matrices were calculated for the estimation of index parameters

Provenance	1	2	3	4	5	6	7	8	9	10	11	12
Family	5	4	4	4	4	4	4	4	5	4	4	3

The C and V matrices in the next section have been calculated from the number of trees in the first family in the sixth provenance. This family has been choosen so that subsequent values of selection index scores can be checked against values in the main text. In addition calculating index parameter estimates from a different family than those in Table AVII-11 will demonstrate that the parameters are relatively insensitive to changes in family sizes for the data set used in the analysis.

i. Calculation of the elements of the C matrix

The calculation is the same as for the homogeneous model except the family variance for provenance 6 is taken from Table AVII-7. The completed C matrix is:

$$\begin{array}{rrr}
0.042 & 0.025 & -0.004 \\
0.025 & 0.034 & -0.015 \\
-0.004 & -0.015 & 0.052 \\
0.133 & 0.090 & -0.030 \\
0.090 & 0.064 & -0.038 \\
-0.030 & -0.038 & 0.171 \\
0.363 & 0.245 & -0.083 \\
0.245 & 0.176 & -0.105 \\
-0.083 & -0.105 & 0.468
\end{array}$$

ii. Calculation of the elements of the V matrix

The calculation is the same as for the homogeneous model except the family and within-family variance components are taken from Tables AVII-7 and AVII-8. The completed V matrix is:

$$\begin{array}{rrrrrrrrr}
0.039 & 0.025 & -0.005 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.025 & 0.040 & -0.017 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
-0.005 & -0.017 & 0.054 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.138 & 0.101 & -0.037 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.101 & 0.095 & -0.050 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & -0.037 & -0.050 & 0.200 & 0.000 & 0.000 & 0.000 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.382 & 0.376 & -0.163 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.376 & 0.566 & -0.248 \\
0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & -0.163 & -0.248 & 0.800
\end{array}$$

V. Calculation of selection index values

Selection index values are calculated by substituting the appropriate values in formula III-2. In Chapter three it is assumed that the economic values are all 1.0. The **C** and **V** matrices have been calculated above for the family 1 in provenance 6. The vector of phenotypic predictors is composed of three sub-vectors:

1. The vector of provenance phenotypic predictors.
2. The vector of family phenotypic predictors.
3. The vector of within-family phenotypic predictors.

The vector of provenance phenotypic predictors for provenance 6 is given in Table AVII-1. The vector of family phenotypic predictors for family 1 in provenance 6 is given in Table AVII-2. The vector of within-family phenotypic predictors for tree 10 in family 1 in provenance 6 is given in Table AVII-3. In summary the **p** vector is:

$$\begin{array}{r} -0.147 \\ -0.253 \\ 0.253 \\ 0.028 \\ -0.141 \\ 0.669 \\ 0.585 \\ -0.569 \\ 0.223 \end{array}$$

Substituting the economic values, the **C** matrix for the homogeneous family variance model, the **V** matrix for the homogeneous family variance model and the **p** vector in formula III-2 gives a homogeneous selection index score of 0.499. This value agrees with homogeneous selection index value for the tree in family 1 in provenance 6 given in Table III-4.

Substituting the economic values, the C matrix for the heterogeneous family variance model, the V matrix for the heterogeneous family variance model and the **p** vector in formula III-2 gives a heterogeneous selection index score of 2.44. This value agrees with heterogeneous selection index value for the tree in family 1 in provenance 6 given in Table III-4.

VI. Calculation of variance of predicted genetic worth, variance of genetic worth and efficiency of the selection index

The calculation of the variance of the genetic worth requires that the G matrix is equal to four times the family variance-covariance matrix plus the provenance variance-covariance matrix. The G for the homogeneous family variance model is calculated from the values given in Tables AVII-4 and AVII-5. The G matrix for the homogeneous family variance model is:

	Height	Circumference	Lean
Height	0.502	0.403	-0.117
Circumference	0.403	0.379	-0.098
Lean	-0.117	-0.098	0.103

The G matrix for the heterogeneous family variance model is calculated from values given in Tables AVII-4 and AVII-7. The G matrix for the heterogeneous family variance model for provenance 6 is:

	Height	Circumference	Lean
Height	0.538	0.359	-0.117
Circumference	0.359	0.275	-0.158
Lean	-0.117	-0.158	0.691

A. Variance of predicted genetic worth

The variance of the predicted genetic worth is calculated by substituting the appropriate values in formula III-12. The C and V matrices for the homogeneous family variance model have been calculated previously. The a vector is assumed to be a vector of ones. The variance of the predicted genetic worth for the homogeneous family variance model is calculated to be 0.826. The variance of the predicted genetic worth for the heterogeneous family variance model using the C and V matrices calculated for provenance 6 is 1.659. This value agrees with the value in Table III-5.

B. Variance of genetic worth

The variance of the genetic worth is calculated by substituting the appropriate values in formula III-13. The G matrix for the homogeneous family variance model has been calculated previously. The variance of the genetic worth for the homogeneous family variance model is calculated to be 1.360. The variance of the genetic worth for the heterogeneous family variance model using the G matrix calculated for provenance 6 is 1.672. This value agrees with the value in Table III-5.

C. Efficiency of selection index

The efficiency is calculated by substituting the appropriate values into formula III-13. The $a'C'V^{-1}Ca$ and a'Ga values have been calculated previously for the homogeneous family variance model and for provenance 6 for the heterogeneous family variance model. Substituting the homogeneous family variance model values into formula III-13 gives:

$$0.826^{0.5} \times 1.36^{-0.5}$$

$$= 0.779$$

Substituting the heterogeneous family variance model values for provenance 6 into formula III-13 gives:

$$1.659^{0.5} \times 1.672^{-0.5}$$

$$= 0.996$$

This value agrees with the value in Table III-5.

VII. Conclusion

The preceding sections of this appendix have demonstrated how the basic input values fit into the formulae given in the main text and Appendix IV. Calculated values agree with those values given in the main text. If the reader desires to check the calculations it is important to start at the beginning of the calculation and go through each step in the computer. If matrices which are produced at various stages are taken out of the computer and then read in rounding and truncation errors will give values which do not agree with those in the main text.

APPENDIX VIII.

NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER FOUR.

I. Introduction

The numerical values which are calculated in Chapter four are presented in the form of response surfaces. The values on the response surfaces represent index parameters. Formulae for calculating each of these parameters are given in Table IV-4. There are two basic elememts in these formulae i.e. $a'C'V^{-1}Ca$ and $a'Ga$. In Chapter four it is assumed that the family variance is homogeneous so the values calculated for $a'C'V^{-1}Ca$ and $a'Ga$ in the homogeneous family variance model in Chapter three can be used to verify calculations in Chapter four. The values calculated in Appendix VII for the homogeneous family variance model are:

$$a'C'V^{-1}Ca = 0.826$$

$$a'Ga = 1.360$$

The a vector in the calculation above is assumed to be a vector of ones. In Chapter four the a vector is changed and the parameter values are recaculated. The response surface represents the parameter values for a range of a values.

II. Values plotted in Chapter four

The values plotted in Chapter four are given in Tables AVIII-1, AVIII-2, AVIII-3, AVIII-4, AVIII-5 and AVIII-6.

Table AVIII-1. Prediction variance response surface values

Circumference economic values	Lean economic values										
	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	9.377	10.598	12.076	13.811	15.803	19.272	20.557	23.320	26.339	29.616	33.149
-8	4.848	5.532	6.472	7.669	9.123	11.786	12.802	15.027	17.508	20.247	23.243
-6	2.157	2.302	2.705	3.364	4.280	6.136	6.883	8.570	10.514	12.715	15.173
-4	1.302	0.909	0.774	0.895	1.274	2.323	2.801	3.950	5.357	7.020	8.939
-2	2.283	1.353	0.680	0.263	0.104	0.347	0.556	1.167	2.036	3.161	4.543
1	7.199	5.462	3.982	2.759	1.793	0.626	0.632	0.436	0.498	0.816	1.392
2	9.756	7.750	6.002	4.510	3.275	1.904	1.575	1.111	0.904	0.953	1.260
4	16.248	13.704	11.417	9.388	7.615	5.437	4.840	3.838	3.093	2.604	2.373
6	24.576	21.494	18.670	16.102	13.792	10.807	9.941	8.401	7.118	6.092	5.323
8	34.741	31.121	27.759	24.654	21.805	18.014	16.879	14.801	12.981	11.417	10.110
10	46.742	42.585	38.685	35.042	31.655	27.058	25.654	23.038	20.679	18.578	16.733

Table AVIII-2. Genetc worth variance response surface values

Circumference economic values	Lean economic values										
	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	23.382	23.126	23.694	25.086	27.302	32.171	34.206	38.894	44.406	50.742	57.902
-8	15.270	14.230	14.014	14.622	16.054	19.747	21.390	25.294	30.022	35.574	41.950
-6	10.190	8.366	7.366	7.190	7.838	10.355	11.606	14.726	18.670	23.438	29.030
-4	8.142	5.534	3.750	2.790	2.654	3.995	4.854	7.190	10.350	14.334	19.142
-2	9.126	5.734	3.166	1.422	0.502	0.667	1.134	2.686	5.062	8.262	12.286
1	16.287	11.719	7.975	5.055	2.959	1.360	1.239	1.615	2.815	4.839	7.687
2	20.190	15.230	11.094	7.782	5.294	3.107	2.790	2.774	3.582	5.214	7.670
4	30.270	24.526	19.606	15.510	12.238	8.875	8.166	7.366	7.390	8.238	9.910
6	43.382	36.854	31.150	26.270	22.214	17.675	16.574	14.990	14.230	14.294	15.182
8	59.526	52.214	45.726	40.062	35.222	29.507	28.014	25.646	24.102	23.382	23.486
10	78.702	70.606	63.334	56.886	51.262	44.371	42.486	39.334	37.005	35.502	34.822

Table AVIII-3. Prediction standard deviation response surface values

Circumference economic values	Lean economic values										
	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	3.062	3.255	3.475	3.716	3.975	4.390	4.534	4.829	5.132	5.442	5.758
-8	2.202	2.352	2.544	2.769	3.020	3.433	3.578	3.876	4.184	4.500	4.821
-6	1.469	1.517	1.645	1.834	2.069	2.477	2.624	2.927	3.243	3.566	3.895
-4	1.141	0.954	0.880	0.946	1.129	1.524	1.674	1.988	2.314	2.649	2.990
-2	1.511	1.163	0.824	0.513	0.322	0.589	0.746	1.080	1.427	1.778	2.131
1	2.683	2.337	1.996	1.661	1.339	0.909	0.795	0.661	0.706	0.904	1.180
2	3.123	2.784	2.450	2.124	1.810	1.380	1.255	1.054	0.951	0.976	1.122
4	4.031	3.702	3.379	3.064	2.759	2.332	2.200	1.959	1.759	1.614	1.540
6	4.957	4.636	4.321	4.013	3.714	3.287	3.153	2.898	2.668	2.468	2.307
8	5.894	5.579	5.269	4.965	4.670	4.244	4.108	3.847	3.603	3.379	3.180
10	6.837	6.526	6.220	5.920	5.626	5.202	5.065	4.800	4.547	4.310	4.091

Table AVIII-4. Genetic worth standard deviation response surface values

Circumference economic values	Lean economic values										
	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	4.835	4.809	4.868	5.009	5.225	5.672	5.849	6.237	6.664	7.123	7.609
-8	3.908	3.772	3.744	3.824	4.007	4.444	4.625	5.029	5.479	5.964	6.477
-6	3.192	2.892	2.714	2.681	2.800	3.218	3.407	3.837	4.321	4.841	5.388
-4	2.853	2.352	1.936	1.670	1.629	1.999	2.203	2.681	3.217	3.786	4.375
-2	3.021	2.395	1.779	1.192	0.709	0.817	1.065	1.639	2.250	2.874	3.505
1	4.036	3.423	2.824	2.248	1.720	1.166	1.113	1.271	1.678	2.200	2.773
2	4.493	3.903	3.331	2.790	2.301	1.763	1.670	1.666	1.893	2.283	2.769
4	5.502	4.952	4.428	3.938	3.498	2.979	2.858	2.714	2.718	2.870	3.148
6	6.587	6.071	5.581	5.125	4.713	4.204	4.071	3.872	3.772	3.781	3.896
8	7.715	7.226	6.762	6.329	5.935	5.432	5.293	5.064	4.909	4.835	4.846
10	8.871	8.403	7.958	7.542	7.160	6.661	6.518	6.272	6.083	5.958	5.901

Table AVIII-5. Prediction error variance response surface values

Circumference economic	Lean economic values										
values	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	14.005	12.528	11.618	11.275	11.499	12.899	13.649	15.574	18.067	21.126	24.753
-8	10.422	8.698	7.542	6.953	6.931	7.961	8.588	10.267	12.514	15.327	18.707
-6	8.033	6.064	4.662	3.826	3.558	4.219	4.723	6.156	8.156	10.723	13.857
-4	6.841	4.625	2.976	1.895	1.380	1.672	2.053	3.240	4.994	7.315	10.203
-2	6.843	4.381	2.486	1.159	0.398	0.320	0.578	1.519	3.026	5.101	7.743
1	9.088	6.257	3.993	2.296	1.166	0.534	0.607	1.179	2.317	4.023	6.295
2	10.434	7.480	5.093	3.272	2.020	1.203	1.215	1.663	2.678	4.261	6.410
4	14.022	10.822	8.189	6.122	4.623	3.438	3.326	3.528	4.297	5.634	7.537
6	18.806	15.360	12.480	10.168	8.422	6.868	6.633	6.589	7.112	8.202	9.855
8	24.785	21.093	17.967	15.408	13.417	11.493	11.135	10.845	11.121	11.965	13.376
10	31.960	28.021	24.649	21.844	19.607	17.313	16.832	16.296	16.326	16.924	18.089

Table AVIII-6. Efficiency response surface values

Circumference economic	Lean economic values										
values	-10	-8	-6	-4	-2	1	2	4	6	8	10
-10	0.633	0.677	0.714	0.742	0.761	0.774	0.775	0.774	0.770	0.764	0.757
-8	0.563	0.623	0.680	0.724	0.754	0.773	0.774	0.771	0.764	0.754	0.744
-6	0.460	0.525	0.606	0.684	0.739	0.770	0.770	0.763	0.750	0.737	0.723
-4	0.400	0.405	0.454	0.566	0.693	0.763	0.760	0.741	0.719	0.700	0.683
-2	0.500	0.486	0.463	0.430	0.455	0.721	0.700	0.659	0.634	0.619	0.608
1	0.665	0.683	0.707	0.739	0.778	0.779	0.714	0.520	0.421	0.411	0.425
2	0.695	0.713	0.736	0.761	0.786	0.783	0.751	0.633	0.502	0.428	0.405
4	0.733	0.748	0.763	0.778	0.789	0.783	0.770	0.722	0.647	0.562	0.489
6	0.753	0.764	0.774	0.783	0.788	0.782	0.774	0.749	0.707	0.653	0.592
8	0.764	0.772	0.779	0.784	0.787	0.781	0.776	0.760	0.734	0.699	0.656
10	0.771	0.777	0.782	0.785	0.786	0.781	0.777	0.765	0.748	0.723	0.693

III. Verification of plotted values

The plotted values will be verified using economic values of 1.0 for each trait. The predicted genetic and genetic worth variances have been calculated previously. The calculated values agree with the values in Tables AVIII-1 and AVIII-2. The response surface values for the prediction standard deviation and the genetic standard deviation are the square root of the variances. The calculated values are:

$$(a'C'V^{-1}Ca)^{0.5} = 0.909$$

$$(a'Ga)^{0.5} = 1.166$$

These values agree with those in Tables AVIII-3 and AVIII-4. The prediction error variance is the difference between the genetic worth variance and the predicted genetic worth variance. The calculated value is 0.534. This value agrees with the value in Table AVIII-5. The efficiency has been calculated in Appendix VII the value is 0.779. This value agrees with the value in Table AVIII-6.

IV. Conclusion

The values plotted in Chapter four have been verified for one set of economic values. The reader may verify the plotted values for other sets of economic values. To obtain the same results as in the tables in this appendix the basic input values given in Appendix VII should be used. Potential errors from truncation should be avoided when calculating values.

APPENDIX IX.

INPUT VALUES AND NUMERICAL EXAMPLES OF CALCULATIONS IN CHAPTER FIVE.

I. Input values

The basic input values for the calculations in chapter V are the are the variance-covariance matrices, the number of trees in each family family and the economic weightings. The calculation of the minmax and Bayes strategies in chapter V are based on a matrix of 121 × 121. This matrix is difficult to represent on a page so a reduced subset of economic weightings has been choosen to demonstrate the techniques.

A. Variance-covariance matrices

The male and female family variance-covariance matrices are given in Tables V-2 and V-3. The within-family variance-covariance matrix is given in Table AIX-1.

Table AIX-1. Within-family variance-covariance matrix

	Minimum density	Mean density	Maximum density
Minimum density	294.27	243.07	192.07
Mean density	243.07	280.29	314.80
Maximum density	192.07	314.80	435.77

B. The number of trees in each family

The number of trees in each family is given in Table AIX-2.

Table AIX-2. Number of trees in each family

Male Family	Female Family								
	1	2	3	4	5	6	7	8	9
1	15	11	10	11	12	12	13	14	
2	7	13	13	14	13	14	15	13	
3	14	10	14	13	11	13	14	11	14
4	16	15	15	13	12	12	12	6	13
5	11	12	15	14	15	14			
6	15	14	13	14	15	14	12	11	12
7	11	9	14	15	12	11	14	11	
8	10	13	8	15	14	13	15	13	
9	13	13	12	14	15	12	15	13	12
10	12	15	15	15	9	13	13		
11	14	15	15	12	13	9	15	13	
12	13	8	15	15	14	14	14	10	
13	14	14	11	14	13	13	14	14	
14	12	13	11	14	12	14	12	13	
15	11	11	7	14	14	12	13	16	
16	13	12	12	10	13	14	15		
17	12	12	15	13	13	13	11	15	
18	13	12	15	13	13	13	11	15	
19	13	15	11	15	13	13	10		
20	13	15							
21	11	13	15	11	14	13	13		
22	12	15	12	12	14	11	10	12	11
23	13	14	15	14	15	13	12	8	
24	10	9	11	13	15	15	14	9	
25	13	15	14	13	13	12	14	15	
26	15	14	13	14	15	15	14	15	12
27	14	13	8	15	9	15	11	13	
28	11	11	14	15	12	12	13	13	
29	10	11	13	13	15	15	11	13	14
30	11	12	12	12	14	14	13	9	12

C. Economic weigthings

The economic weigthings used in chapter V are given in Tables V-5 and V-6. The economic weightings used in the numerical examples in this appendix are given in Table AIX-3.

Table AIX-3. Set of economic weights in numerical example (minimum density held constant).

Mean density	-10	-5	1	5	10
Maximum density	-10	-5	1	5	10

II. Calculation of gain

Gain is calculated by substituting tje correct values in formula V-26. The C and V matrices vary from family to family the example in this appendix will used the fourth female family nested within the thirteenth male family. The values vary because of the number of trees in the male and female family. The values for the number of trees in each family can be read off Table AIX-2. There are two independant genetic values to be predicted. There are therefore two $C'V^{-1}C$ matrices which can be added together to give the combined $C'V^{-1}C$ matrix. The C matrices are different for the prediction of additive genetic values of individual trees due to the male parent contribution and for the prediction of the additive genetic value of the individual tree due to the female parent contribution. The V^{-1} matrix is the same for each predited additive value.

A. C matrix for prediction of the additive genetic value of the individual due to the male parent contribution

There are three sub-matrices which correspond to the covariance between the genetic values and:

1. The male family phenotypic predictors
2. The female family phenotypic predictors
3. The within-family phenotypic predictors

i. The C sub-matrix between the genetic values and the male family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-13. Substituting the number of trees in the female and male families in the formula gives:

$$((.5 + ((107. - 1.) \times .25)) + 107.) \times \sigma_{AM}^{2}$$

The paternal genetic variance (σ_{AM}^{2}) is given by multiplying the male family variance-covariance by 4. For example the minimum density paternal additive genetic variance would be 4 times 31.03 (Table V-2). Substituting this value in formula V-13 gives:

$$((.5 + ((107. - 1.) \times .25)) + 107.) \times 124.12$$

$$= 31.32$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate paternal additive genetic variance or covariance replaces the paternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	31.32	30.05	27.91
Mean density	30.05	28.72	26.53
Maximum density	27.91	26.53	24.32

ii. The C sub-matrix between the genetic values and the female family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-14. Substituting the number of trees in the female and male families in the formula gives:

$$\begin{gathered}(((.5 + ((14. - 1.) \times .25)) \div 14.) \\ - ((.5 + ((107. - 1.) \times .25)) \div 107.)) \times \sigma_{AM}^{2}\end{gathered}$$

Substituting the paternal additive genetic variance for minimum density gives:

$$\begin{gathered}(((.5 + ((14. - 1.) \times .25)) \div 14.) \\ - ((.5 + ((107. - 1.) \times .25)) \div 107.)) \times 124.12 \\ = 1.93\end{gathered}$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate paternal additive genetic variance or covariance replaces the paternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	1.93	1.85	1.72
Mean density	1.85	1.77	1.63
Maximum density	1.72	1.63	1.50

iii. The C sub-matrix between the genetic values and the within-family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-15. Substituting the number of trees in the female family in the formula gives:

$$(.5 - ((.5 + ((14. - 1.) \times .25)) \div 14.)) \times \sigma_{AM}^{2}$$

Substituting the paternal additive genetic variance for minimum density gives:

$$(.5 - ((.5 + ((14. - 1.) \times .25)) \div 14.)) \times 124.12$$

$$= 28.81$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate paternal additive genetic variance or covariance replaces the paternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	28.81	27.64	25.67
Mean density	27.64	26.42	24.40
Maximum density	25.67	24.40	22.30

In summary the completed C' matrix is:

$$\begin{matrix} 31.32 & 30.05 & 27.91 & 1.93 & 1.85 & 1.72 & 28.81 & 27.64 & 25.67 \\ 30.05 & 28.72 & 26.53 & 1.85 & 1.77 & 1.63 & 27.64 & 26.42 & 24.40 \\ 27.91 & 26.53 & 24.32 & 1.72 & 1.63 & 1.50 & 25.67 & 24.40 & 22.38 \end{matrix}$$

B. C matrix for prediction of the additive genetic value of the individual due to the female parent contribution

There are three sub-matrices which correspond to the covariance between the genetic values and:

1. The male family phenotypic predictors
2. The female family phenotypic predictors
3. The within-family phenotypic predictors

i. The C sub-matrix between the genetic values and the male family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-8. Substituting the number of trees in the female and male families in the formula gives:

$$((.5 + ((14. - 1.) \times .25)) \div 107.) \times \sigma_{AF}^{2}$$

The maternal genetic variance (σ_{AF}^{2}) is given by multiplying the female family variance-covariance by 4. For example the minimum density maternal additive genetic variance would be

4 times 17.47 (Table V-3). Substituting this value in formula V-8 gives:

$$((.5 + ((14. - 1.) \times .25)) + 107.) \times 69.88$$

$$= 2.45$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate paternal additive genetic variance or covariance replaces the maternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	2.45	2.41	2.29
Mean density	2.41	2.33	2.16
Maximum density	2.29	2.16	1.94

ii. The C sub-matrix between the genetic values and the female family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-11. Substituting the number of trees in the female and male families in the formula gives:

$$(((.5 + ((14. - 1.) \times .25)) + 14.)$$
$$- ((.5 + ((14. - 1.) \times .25)) + 107.)) \times \sigma_{AF}^{2}$$

Substituting the maternal additive genetic variance for minimum density gives:

$$(((.5 + ((14. - 1.) \times .25)) \div 14.) - ((.5 + ((14. - 1.) \times .25)) \div 107.)) \times 68.88$$

$$= 16.27$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate maternal additive genetic variance or covariance replaces the paternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	16.27	15.99	15.19
Mean density	15.99	15.45	14.33
Maximum density	15.19	14.33	12.88

iii. The C sub-matrix between the genetic values and the within-family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-12. Substituting the number of trees in the female family in the formula gives:

$$(.5 - ((.5 + ((14. - 1.) \times .25)) \div 14.)) \times \sigma_{AF}^{2}$$

Substituting the maternal additive genetic variance for minimum density gives:

$$(.5 - ((.5 + ((14. - 1.) \times .25)) \div 14.)) \times 69.88$$

$$= 16.22$$

The caluclation of the other elements of the sub-matrix is the same except the appropriate paternal additive genetic variance or covariance replaces the paternal additive genetic variance for minimum density. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	16.22	15.94	15.15
Mean density	15.94	15.40	14.29
Maximum density	15.15	14.29	12.84

In summary the completed C' matrix is:

2.45	2.41	2.29	16.27	15.99	15.19	16.22	15.94	15.15
2.41	2.33	2.16	15.99	15.45	14.33	15.94	15.40	14.29
2.29	2.16	1.94	15.19	14.33	12.88	15.15	14.29	12.84

C. V matrix for prediction of additive genetic values

There are three sub-matrices which correspond to:

1. The variance of the male family phenotypic predictors.
2. The variance of the female family phenotypic predictors.
3. The variance of the within-family phenotypic predictors.

i. The V sub-matrix of the male family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-16.

Substituting the number of trees in the male and female families in the formula gives:

$$(((11449. \times \sigma_{M}^{2}) + (1439. \times \sigma_{F(M)}^{2})) + 11449.) + ((1. + 107.) \times \sigma_{e}^{2}$$

Sustituting the variance component values for minimum density (Tables V-2, V-3 and AIX-1) gives:

$$(((11499. \times 31.03) + (1439 \times 17.47)) + 11449) + ((1. + 107.) \times 294.27)$$

$$= 35.98$$

The calculation of the other elements of the sub-matrix is the same except the appropriate variances and covariances replace the minimum density variance components. The completed submatrix is:

	Minimum density	Mean density	Maximum density
Minimum density	35.98	34.20	31.50
Mean density	34.20	33.15	31.16
Maximum density	31.50	31.16	29.91

ii The V sub-matrix of female family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-17. Substituting the number of trees in the male and female families gives:

$$(((11449. - 1439.) + 11449.) \times \sigma_{F(M)}^{2})$$
$$+ (((1. + 14.) - (1. + 107.)) \times \sigma_{e}^{2})$$

Substituting the variance component value for minimum density gives:

$$(((11449. - 1439.) + 11449.) \times 17.47)$$
$$+ (((1. + 14.) - (1. + 107.)) \times 294.27)$$
$$= 33.54$$

The calculation of the other elements of the sub-matrix is the same except the appropriate variance or covariance components replace the minimum density variance components. The completed sub-matrix is:

	Minimum density	Mean density	Maximum density
Minimum density	33.54	30.10	26.18
Mean density	30.10	31.91	33.00
Maximum density	26.18	33.00	39.14

ii The V sub-matrix of within-family phenotypic predictors

The individual elements of the sub-matrix are given by substituting the correct values into formula V-18. Substituting the number of trees in the female family gives:

$$((14. - 1.) \div 14.) \times \sigma_{e}^{2}$$

Substituting the variance component value for minimum density gives:

$$((14. - 1.) \div 14.) \times 294.27$$

$$= 273.25$$

The calculation of the other elements of the sub-matrix is the same except the appropriate variance or covariance component replaces the minimum density variance component. The complete submatrix is:

	Minimum density	Mean density	Maximum density
Minimum density	273.25	225.71	178.35
Mean density	225.71	260.27	292.31
Maximum density	178.35	292.31	404.64

In summary the completed V sub-matrix is:

35.98	34.20	31.50	0.00	0.00	0.00	0.00	0.00	0.00
34.20	33.15	31.16	0.00	0.00	0.00	0.00	0.00	0.00
31.50	31.16	29.91	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	33.54	30.10	26.18	0.00	0.00	0.00
0.00	0.00	0.00	30.10	31.91	33.00	0.00	0.00	0.00
0.00	0.00	0.00	26.18	33.00	39.14	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	273.25	225.71	178.35
0.00	0.00	0.00	0.00	0.00	0.00	225.71	260.27	292.31
0.00	0.00	0.00	0.00	0.00	0.00	178.35	292.31	404.64

D. Calculation of $C'V^{-1}C$ matrices

There are two $C'V^{-1}C$ matrices:

1. The $C'V^{-1}C$ matrix associated with the additive genetic value due to the male parent contribution.
2. The $C'V^{-1}C$ matrix associated with the additive genetic value due to the female parent contribution.

Using the C and V matrices calculated previously and substituting the values into $C'V^{-1}C$ gives the matrix for the male parent contribution:

	Minimum density	Mean density	Maximum density
Minimum density	30.89	29.62	27.47
Mean density	29.62	28.45	26.45
Maximum density	27.47	26.45	24.66

Substituting the C and V values for the female parent contribution in $C'V^{-1}C$ gives:

	Minimum density	Mean density	Maximum density
Minimum density	9.74	9.57	9.04
Mean density	9.57	9.44	8.97
Maximum density	9.04	8.97	8.57

Adding the two matrices gives:

	Minimum density	Mean density	Maximum density
Minimum density	40.64	39.18	36.51
Mean density	39.18	37.89	35.42
Maximum density	36.51	35.42	33.23

E. Substitution of $C'V^{-1}C$ matrix in the gain formula

Substitution of the $C'V^{-1}C$ matrix in formula V-26 will give the expected gain. Selection intensities of 1.4 and 1.96 were choosen in the main text. To simplify the calculations in this appendix a selection intensity of 1.0 will be assumed. Consider the example of economic weightings at the time of selection of:

Mean density = 1.

Minimum density = -10.

Maximum density = -10.

Substituting the values in $a_s'C'V^{-1}Ca_s$ gives 12723.42. Taking the square root of this value gives 112.80 which is the denominator in formula V-26.

Assuming economic weigths at the time of selection to be:

Mean density = 1.
Minimum density = -5.
Maximum density = -5.

Substituting the above values in $a_H'C'V^{-1}Ca_s$ gives 6003.57 which is the value of the numerator in formula V-26. Taking the selection intensity to be 1.0 the expected gain from selection based on the above combination of economic weights at the time of selection and at the time of harvest is:

$$1. \times (6003.57 \div 112.80)$$

III. Calculation of min-max strategy

The expected gain values for the range of economic values considered in this appendix are given in Table AIX-4.

Table AIX-4. Expected gain for ranges of economic values.

Values at harvest time	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
1 -10 -10	112.80	112.80	-112.80	-112.80	-112.80
1 -5 -5	53.23	53.23	-53.23	-53.23	-53.23
1 1 1	-18.28	-18.28	18.28	18.28	18.28
1 5 5	-65.93	-65.93	65.93	65.93	65.93
1 10 10	-125.50	-125.50	125.50	125.50	125.50

The minimum outcome values for each set of economic values at selection time are given in Table AIX-5.

Table AIX-5. Minimum outcome values for each set of economic values at selection time.

	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
Minimum values	-125.50	-125.50	-112.80	-112.80	-112.80

The strategy or economic values at selection time which gives the maximum minimum value in Table AIX-5. is the minmax strategy. In the above example three strategies have the maximum value of -112.80.

IV. Calculation of Bayes strategy

Bayes strategy is based on the gain values in Table AIX-4. The Bayes strategy requires that probabilities be assigned to the economic values at harvest time. There is little or no information on these values so it is assumed as in Chapter five that each set of economic values is equally probable. In this case each set of values has a probability of .2. The value for the strategy value of economic weights of 1 -10 -10 at selection time is:

$$(112 \times .2) + (53.23 \times .2) + (-18.28 \times .2) + (-65.93 \times .2) + (-125.50 \times .2)$$

$$= -8.74$$

The optimum strategy or Bayes strategy is the strategy which has the highest value. In this case there are three strategies which are equal with a value of 8.74.

Table AIX-6. Bayes strategy values for each set of economic values at selection time.

	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
Strategy values	-8.74	-8.74	8.74	8.74	8.74

V. Calculation of minmax regret strategy

The minmax regret strategy is based on similar logic to the minmax strategy except the values by which stragties are assessed are regret values. Regret is expressed as the difference between the highest value outcome given a set of economic weightings at harvest time. For example consider the gains matrix in Table AIX-4. The regret for economic values of 1 1 1 at selection time if the values at harvest time turn out to be 1 -10 -10 is 112.8 - 112.8. The regret is 225.6. Table AIX-7 contains the regret values which have been determined from the gain values in Table AIX-4.

Table AIX-7. Expected regret values for ranges of economic values.

Values at harvest time	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
1 -10 -10	0.00	0.00	225.60	225.60	225.60
1 -5 -5	0.00	0.00	106.46	106.46	106.46
1 1 1	36.56	36.56	0.00	0.00	0.00
1 5 5	131.86	131.86	0.00	0.00	0.00
1 10 10	251.0	251.0	0.00	0.00	0.00

The rational human being is assumed to select the strategy which minimizes regret. Thus when considering regret the minmax strategy will be the one which minimizes the maximum regret. There are three sets of economic weightings which have the minimum maximum regret value of 225.60 (Table AIX-8).

Table AIX-8. Maximum regret values for each set of economic values at selection time.

	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
Maximum values	251.0	251.0	225.6	225.6	225.6

VI. Calculation of Bayes regret strategy

The Bayes regret strategy is calculated in the same way as the Bayes strategy except regret values are used and the best strategy is the one which has the lowest regret values. In this case there are three strategies with the lowest regret value of 66.41 (Table AIX-9).

Table AIX-9. Bayes regret strategy values for each set of economic values at selection time.

	Economic values at selection time				
	1 -10 -10	1 -5 -5	1 1 1	1 5 5	1 10 10
Strategy values	83.88	83.88	66.41	66.41	66.41

VII. Conclusion

The numerical examples above have been calculated on a computer. The intermediate values to the calculation shown above will not give the same answers as the computer because values have been rounded to enable a clear presentation of values. If the reader wishes to verify the calculate values it will be necessary to start with the basic input values and then compair computer output with the rounded values given in this appendix. A note of caution should be added the reader who can work through the examples above does not necessarily understand the calculation. The reader is recommended to the references in Chapter five if he or she wants to understand the mathematics.